

②

ARO-UR Center for

DTIC FILE COPY

AD-A218 425

OPTO-ELECTRONIC SYSTEMS RESEARCH

TECHNICAL REPORT

QUANTUM-LIMITED IMAGE RECOGNITION

Thomas Arthur Isberg

December 1989

The Institute of Optics
University of Rochester

DTIC
ELECTE
FEB 26 1990
S E D

Prepared for:

U.S. Army Research Office
ATTN: DRXRO-IP-Library
P. O. Box 12211
Research Triangle Park, NC 27709

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

90 02 26 194

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			5. MONITORING ORGANIZATION REPORT NUMBER(S) <i>ARO 24626.128-PH-01R</i>		
6a. NAME OF PERFORMING ORGANIZATION University of Rochester		6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION U. S. Army Research Office		
6c. ADDRESS (City, State, and ZIP Code) The Institute of Optics Rochester, New York 14627			7b. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U. S. Army Research Office		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <i>DAAL03-86-K-0173</i>		
8c. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.
			WORK UNIT ACCESSION NO.		
11. TITLE (Include Security Classification) Quantum-Limited Image Recognition					
12. PERSONAL AUTHOR(S) Thomas Arthur Isberg					
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) December 1989	
15. PAGE COUNT 194					
16. SUPPLEMENTARY NOTATION The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD GROUP SUB-GROUP			automatic target recognition; photon counting quantum-limited imaging		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) See Page iv-v					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL G. Michael Morris			22b. TELEPHONE (Include Area Code) 716-275-5140		22c. OFFICE SYMBOL

DTIC
ELECTE
FEB 26 1990
S E D

Quantum-Limited Image Recognition

by

Thomas Arthur Isberg

Submitted in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

Supervised by Professor G. Michael Morris

The Institute of Optics

University of Rochester

Rochester, New York

1989



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Curriculum Vitae

Thomas Isberg was born on September 24, 1961 in Oak Park, Illinois. He began his undergraduate work in 1979 at the University of Illinois at Chicago, where he received a Bachelor of Science degree in Physics in 1983. In the fall of 1983, the author began graduate studies at the University of Rochester.

Acknowledgements

I would like to express my sincere thanks to my advisor, Prof. G. Michael Morris for his guidance, support and patience throughout my studies at the Institute of Optics. I am also grateful to other members of the faculty, staff and graduate students at the Institute. I would especially like to thank all of the members of our research group, particularly Dean, Lenny, Tony and Miles and Lenore.

Special thanks to my wife Helen, for all of her love, support and understanding during her "forced" stay in Rochester. I would also like to thank my Father and Grandfather, who always supported my efforts, but did not get the chance to see their fruition. Finally, I would like to thank the friends I have made during my studies here in Rochester, and it is with much sadness that I must leave them upon completion of this work.

I gratefully acknowledge the support I have received from the New York State Center for Advanced Optical Technology, and the U.S. Army Research Office.

Abstract

Correlation-based methods for automatic image recognition are implemented using a position-sensitive, photon-counting detection system. It is demonstrated that the information provided by a small number of detected photoevents can be used to accurately estimate the cross correlation between a classical-intensity input scene and a reference (or filter) function stored in computer memory. A theoretical formalism is developed that describes the behavior of the quantum-limited correlation signal for complex filter functions. The theoretical predictions are verified experimentally using a position-sensitive photon-counting detection system. The speed at which the detection system operates makes this an effective technique for implementing correlation based methods for image recognition in real time, even when there is an abundance of input illumination.

First, image correlation at low light levels is investigated. When the reference function that is stored in computer memory is a digitized version of the classical-intensity input object, the correlation output corresponds to that of a conventional matched filter. It is demonstrated that as few as 1000 detected photoevents provide sufficient information to discriminate accurately among a set of engraved portrait images.

Rotation-invariant image recognition using a rotation-invariant circular-harmonic filter is also implemented using photon-counting techniques. In addition, a new method is demonstrated for normalizing the correlation output in real time using the positional information from the detected photoevents. This new normalization may allow rotation-invariant circular-harmonic filters to be utilized in a cluttered environment.

The estimation of moment invariants for image recognition is also considered. Experiments are performed that demonstrate that the information provided by a few thousand detected photoevents is sufficient to estimate moment invariants that remain unchanged when segmented input images are scaled, change in position, or undergo in-plane rotations.

Finally, the automatic recognition of images from within a cluttered environment is considered. The photon-counting detection system is used to implement a two-stage template matching algorithm to locate objects of interest from within cluttered scenes. Both two-stage matched filtering, and two-stage rotation-invariant filtering is considered.

Table of Contents

Chapter 1. Introduction	1
1.1 Introduction to Machine Vision	1
1.2 Correlation-Based Methods for Image Recognition	5
1.2.1 Template Matching and Matched Filtering.....	6
1.2.2 Geometrical-Distortion Invariant Filters	8
1.2.3 Classification Filters	9
1.3 Overview of the Thesis.....	10
References for Chapter 1	12
 Chapter 2. Correlation with a Deterministic Reference Function.....	 19
2.1 Introduction.....	19
2.2 Detection of Photon-Limited Images	20
2.3 Correlation with a Complex-Valued Reference.....	21
2.3.1 Poisson-Distributed Number of Detected Photoevents	21
2.3.2 N Fixed.....	31
2.4 Correlation with a Real Reference Function	34
2.4.1 N Poisson Distributed	34
2.4.2 N Fixed.....	35
2.5 Probability of Detection and False Alarm	36
2.6 Summary of Chapter 2	38
References of Chapter 2	39
 Chapter 3. Image Correlation at Low Light Levels.....	 40
3.1 Introduction	40
3.2 Quantum-Limited Imaging Systems	40

3.3 Correlation with a Fixed Number of Detected Photoevents.....	47
3.3.1 Theoretical Predictions.....	47
3.3.2 Experimental Results.....	53
3.4 N Poisson Distributed.....	56
3.5 Effects of Additive Noise.....	61
3.5.1 Introduction.....	61
3.5.2 Effect on recognition performance	62
3.5.3 Minimizing the Effects of Additive Noise	67
3.6 Summary.....	71
References for Chapter 3	72

Chapter 4. Rotation-Invariant Image Recognition at Low Light

Levels	74
4.1 Introduction.....	74
4.2 Rotation-Invariant Filtering using Circular-Harmonic Functions	76
4.3 Rotation-Invariant Filtering at Low Light Levels	79
4.4 Experimental Results	83
4.5 Rotation-Invariant Recognition in a Cluttered Environment	93
4.5.1 Normalization using Photon-Counting Techniques.....	95
4.6 Discussion	101
4.7 Summary.....	102
References for Chapter 4.....	103

Chapter 5. Monte Carlo Estimation of Moment Invariants for	
 Pattern Recognition.....	105
5.1 Introduction.....	105
5.2 Moment Invariants for Pattern Recognition.....	107
5.2.1 Basic definitions.....	107
5.2.2 Radial Moments of Circular-Harmonic Functions	
(CHF's).....	108
5.3 Estimation of Radial Moments of CHF's using Photon-Counting	
Techniques.....	112
5.3.1 Photon-Counting Methods for Image Recognition.....	112
5.3.2 Estimation of the Input Centroid.....	114
5.3.3 Estimation of Moment Invariants.....	116
5.3.3.1 Estimation of $\Phi_{2,m}$	117
5.3.3.2 Estimation of $\Phi_{1,m}$	119
5.3.4 Determining the Required Number of Detected Photoevents	
N.....	119
5.4 Experimental Results	123
5.4.1 Estimation of Moment Invariants.....	125
5.4.2 Pattern Recognition Performance.....	134
5.5 Summary.....	138
References for Chapter 5	139

Chapter 6. Two-Stage Template Matching using	
 Quantum-Limited Images.....	141
6.1 Introduction	141
6.2 Two-Stage Template Matching.....	143
6.4 Two-stage Template Matching using Quantum-Limited Images	146
6.4.1 Derivation of the Quantum-Limited Cost Function.....	147

6.4.2 Determination of Correlation Thresholds and Number of Detected Photoevents	150
6.5 Computer Simulations.....	156
6.6 Two-Stage Invariant Filtering using Circular-Harmonic Functions	166
6.7 Discussion	166
6.8 Summary.....	169
References for Chapter 6.....	170
Chapter 7 Concluding Remarks	172
 Appendix A Statistics of the Photon-Limited Correlation Signal.....	 177
A.1 Statistics of the Photon-Limited Correlation Signal	177
A.2 Statistics of the Distance in Feature Space, D^2	184
A.3 Statistics of the Estimate for the Input Centroid.	187
References for Appendix A	189
Appendix B. Issues in the Computation of Weak Moments.....	190

List of Figures

Figure 2.1 System diagram for quantum-limited image recognition.....	23
Figure 2.2 Probability density functions under hypotheses H_0 and H_1	37
Figure 3.1 Images of engraved portraits obtained using two-dimensional photon-counting detection system.....	43
Figure. 3.2 Schematic diagram of a resistive-anode based photon-counting detection system.	46
Figure 3.3. Probability density functions of the correlation signal when the input image $f(x',y')$ is the portrait of (I) George Washington, (II) Abraham Lincoln, and (III) Andrew Jackson. The number of detected photoevents is (a) $N=250$, (b) $N=500$, and (c) $N=1000$. The reference function $R(x',y')$ in all cases is the portrait of George Washington with $N=20$ million.....	51
Figure 3.4. ROC curves for the portraits of Washington and Lincoln for different values of N . The reference function $R(x',y')$ in all cases is the portrait of George Washington with $N=20$ million	52
Figure 3.5 Histogram of correlation values obtained from laboratory measurements of the photon-limited correlation signal using a fixed number of detected photoevents.....	55
Figure 3.6 Probability density functions of the correlation signal realized using a Poisson distributed number of detected photoevents.	59
Figure 3.7 ROC curves for the portraits of Washington and Lincoln for different values of \bar{N}	60
Figure 3.8 Effect of additive noise on the probability density functions of the correlation signal.	65

Figure 3.9 ROC curves for the portraits of Washington and Lincoln in the presence of various amounts of additive noise.....	66
Figure 3.10 ROC curves for the portraits of Washington and Lincoln in the presence of various amounts of additive noise, with the bias subtracted.	70
Figure 4.1 Schematic diagram of resistive-anode based position-sensitive, photon-counting detection system.....	84
Figure 4.2 Input objects and associated second-order circular harmonics.....	85
Figure 4.3 Histogram of experimental values of the squared modulus of the correlation signal when the input image is: I, vise grips, rotated by 90 deg. with respect to the reference; II, crescent wrench and III, pliers. The number of detected photoevents is $N=3000$	89
Figure 4.4 Histogram of values for the rotation angle α when the vise grips are input rotated by 90 deg. with respect to the input. The number of detected photoevents is $N=3000$	90
Figure 4.5 ROC curves for the vise grips and crescent wrench. The vise grips are rotated by 90 deg. with respect to the input.....	91
Figure 4.6 Comparison of theoretical probability density functions for the squared modulus of the correlation signal when the number of detected photoevents is fixed ($N=3000$) and when the number of detected photoevents is Poisson distributed ($\bar{N}=3000$).	92
Figure 4.7 Rotation-invariant filtering in a cluttered environment.	100
Figure 5.1 Engraved portrait objects used in image recognition experiments.	124
Figure 5.2 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Lincoln. The solid lines indicate theoretical predictions for the probability density functions.....	131

Figure 5.3 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Washington. The solid lines indicate theoretical predictions for the probability density function.....	132
Figure 5.4 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Jackson. The solid lines indicate theoretical predictions for the probability density functions.....	133
Figure 5.5. Histograms of experimentally obtained estimates for the distance in feature space D^2 between the portraits of Lincoln and the portraits of Washington, using the features $\Phi_{2,2}$ and $\Phi_{2,4}$. The solid lines indicate theoretical predictions.....	136
Figure 5.6. Histograms of experimentally obtained estimates for the distance in feature space D^2 between the portraits of Lincoln and the portraits of Jackson using the features $\Phi_{2,2}$ and $\Phi_{2,4}$. The solid lines indicate theoretical predictions.....	137
Figure 6.1. Quantum-limited input scene showing reference function window at offset location (x,y).	144
Figure 6.2. Two-stage template matching in a scene with segmented objects.....	157
Figure 6.3 Two-stage template matching in a satellite image.....	160
Figure 6.4 Cost function for the input and reference functions in Fig. 6.3.....	161
Figure 6.5 Attempt at two-stage template matching in an aerial photograph using a reference function with insufficient resolution.....	163
Figure 6.6 Two-stage template matching in an aerial photograph using a reference function with increased resolution.	164
Figure 6.7 Cost function for the input and reference images in Fig. 6.6.	165
Figure 6.8 Two-stage rotation-invariant filtering in an aerial photograph.	168

List of Tables

Table 3.1 Comparison of theoretical predictions and experimental results for the photon-limited correlation signal realized using various number of detected photoevents. In each case, the reference image was the image of Washington (N=20 million) shown in Fig. 3.1	54
Table 4.1. Comparison of theoretical predictions and experimental results for the expected values and standard deviations of the squared modulus of the correlation signal. The number of detected photoevents is N=3000.....	87
Table 5.1. Comparison of high-light-level moments and photon-limited estimates for $\Phi_{1,m}$, m=1 to 9. The input image is Lincoln.	128
Table 5.2. Comparison of high-light-level moments and photon-limited estimates for $\Phi_{2,m}$, m=1 to 9. The input image is Lincoln.	129
Table 5.3. Comparison of high-light-level moments and photon-limited estimates for $\Phi_{3,m}$, m=1 to 9. The input image is Lincoln.	130

Introduction

1.1. Introduction to Machine Vision

Machine vision is a subject that has received much attention in recent years. Several books^{1,2}, journals^{3,4}, and 15,384 journal articles⁵ have been written on various areas of the subject. Some examples of the types of tasks that fall under the broad heading of machine vision include image or pattern recognition, detection of an object's motion, the measurement of an object's surface height profile or texture, edge detection or metrology, and object tracking. Each of the above tasks has found numerous applications in a wide variety of fields. For example, optical-digital hybrid techniques for non-contact surface profilometry have been used in applications such as alignment of parts in automated welding⁶ and the inspection of printed circuit boards⁷. Image recognition tasks range from the inspection of objects for quality control in an industrial environment to the automatic recognition and tracking of military targets in "smart" weapons systems.

In particular, the area of image recognition has received a large amount of attention in recent years. While much has been written on the subject, a "fool proof" method to automatically identify an object in real time, independent of the environment it is placed in, and independent of variations such as rotation, scale, aspect, illumination, and clutter has not been developed to date. Rather, the work in the area of image recognition has achieved its greatest success by concentrating on a particular problem or application, and employing a method or combination of methods that yield acceptable results for that particular problem.

The approaches to the image recognition problem are often placed into three broad categories; namely artificial intelligence approaches¹, structural approaches⁸ and

statistical approaches⁹. The so-called statistical approaches include methods that place an input object into a finite number of classes. Using training sets consisting of different variations that may be typical for a class of objects of interest, statistical techniques are often used to generate correlation filters. The output from the correlation filters is then used to classify the input objects. The classes can be restricted to include variations of a single object, in which case the problem is referred to as image recognition. In other applications, the classes may be taken to be much broader, such as dogs and cats, or trucks and jeeps. All linear filtering, or correlation-based methods for image recognition are usually placed under the heading of statistical approaches. Both optical¹⁰⁻²³ and digital²⁴⁻⁶³ implementations of these correlation-based methods for image recognition have been reported. The parallelism inherent in the optical methods yields a distinct advantage in terms of speed of implementation, but certain characteristics of the optical systems may make them impractical for many applications. For example, the physical size of the optical systems that implement the recognition algorithms, or requirements for the input scene illumination may not be acceptable in many applications. The digital implementations retain the advantage of being much more flexible than the optical implementations, but are sometimes difficult to perform in real time. The large number of computations involved in the digital implementations of the various methods for image recognition are what inhibit their real-time performance.

The limitations of both the optical and digital systems have led researchers to consider novel optical-digital hybrid systems, and novel architectures that combine the flexibility of digital systems with the speed of optical systems⁶⁴⁻⁷⁸. In addition, much work has been performed involving optical correlations using white light^{12,19-22}, which eliminates the need for a coherent input to the optical system.

The artificial intelligence (AI)^{1,83} and the structural⁸ approaches to image recognition are perhaps the most sophisticated, as well as the most computationally intensive techniques. The AI approach uses the description of abstract concepts and the recognition of instances in the input scenes. Typically, hypotheses and inferences are implemented as IF-THEN rules, from which recognition decisions regarding the input scene are deduced⁸³.

In the structural approach, the input scene is considered to be composed of a set of vision primitives, which are arranged in a natural hierarchy^{78,79}. The vision primitives typically consist of allowable shapes that describe the geometrical structure of the input scene. For example, the geometric structure of the input scene has been described in a polyhedral representation that describes explicitly the object's faces, vertices, etc⁷⁸. Objects that may appear at arbitrary orientations and locations in the scene are identified by finding transformation parameters that specify changes in orientation and translation between parts of the scene and object models⁷⁸. The recognition of objects then requires the following three operations. First, one must build a polyhedral representation of the scene and object models in terms of planes, edges, and vertices. Next, one must find the orientation of objects with respect to the object models; and finally, one must find their relative translations with respect to the appropriately oriented object models⁷⁸. Clearly, performing the necessary computations in the implementation of this process is difficult to perform in real time. Recently, much progress has been made in the computation of low-level image analysis using parallel computer architectures such as the commercially available Butterfly Parallel Processor⁸⁰.

Other novel approaches to the image recognition process have received much attention recently. The concept of sensor fusion⁸¹⁻⁸² in performing image recognition

tasks has been demonstrated to be useful in various applications. Sensor fusion involves the use of information provided by multiple sensors about an input object. Such data may include information about an object's thermal signature or other spectral information, or its motion, color, etc. In addition, it has been recently proposed⁷⁹ that ideas connected with animal vision, will play an increasingly important role in shaping computer-vision research. Finally, the use of neural networks⁸⁴ for image recognition^{85,86} has shown promising results in certain applications.

In recent years, a large amount of effort has been directed at the development of methods for recognizing images independent of variations that may occur when the images are input to the recognition system. The variations in the input are usually placed into two categories, namely intrinsic variations and geometric distortions. Intrinsic variations include variations within a class of objects, or changes resulting from non-uniform illumination, or occlusions of the object. Geometrical distortions include changes in the object's orientation, scale, or position. Structural, AI and correlation based approaches have all been applied to this problem with varying degrees of success, depending upon the application.

The approach to image recognition detailed in this thesis involves a novel implementation of several of the correlation-based techniques for recognizing images independent of geometric distortions⁷⁰⁻⁷⁸. In these methods, the information that is contained in the input scene is statistically sampled by reducing the input illumination to very low light levels, and imaging the resultant quantum-limited input scene onto a position-sensitive, photon-counting detection system⁸⁷⁻⁸⁸. The probabilistic relationship between the spatial coordinates of the detected photoevents and the classical intensity of the corresponding location in the input scene⁸⁹ provides a natural means of sampling the information contained in the input scene. When the photon-limited input

scene is cross-correlated with a filter function (derived from one of the statistical approaches for image recognition), the resulting quantum-limited correlation signal provides an estimate for the high-light-level correlation signal. In essence, the method may be regarded as an optical method to implement a Monte Carlo⁹⁰ scheme to estimate correlation integrals. Depending upon the application, the number of detected photoevents required to accurately estimate the high-light-level cross-correlation signal may range between a few hundred⁷⁷ to a few thousand⁷³. The detection system employed can detect photoevents at rates up to 100 kHz., and the correlation signal can be computed in real-time. As a result, the total computation time can be on the order of tens of milliseconds for many images. The detection system is intensity based, and the correlations between the photon-limited input scene and a reference function stored in computer memory are performed digitally. Hence, this optical-digital hybrid system retains the advantages of the digital implementations of the correlation-based methods for image recognition, while providing for relatively short computation times.

In this chapter, some of the most pertinent correlation-based methods for image recognition are briefly reviewed, and compared. In addition, the reasoning behind the selection of the particular methods employed in this thesis is explained. Finally, a brief overview of the work presented in this thesis is given.

1.2 Correlation-Based Methods for Image Recognition

1.2.1 Template Matching and Matched Filtering

Correlation-based methods were one of the first techniques applied to automatic image recognition⁵². The most obvious implementation of a correlation-based method is template matching. In template matching, the input and reference (template) images are considered to be vectors. One employs some metric (such as a Euclidean norm, or

normalized cross-correlation) to determine the similarity between the input object and the template. If the similarity measure exceeds some threshold, the object is determined to match the template. Conversely, if the similarity measure does not exceed the threshold, the object is said to be different from the template. The motivation for using the normalized cross-correlation as the metric is provided by the Schwarz inequality⁹¹. The Schwarz inequality states the following about the cross correlation between two (in general complex) functions $f(x',y')$ and $g(x',y')$:

$$\left| \int \int f(x',y') g^*(x',y') dx' dy' \right|^2 \leq \left(\int \int |f(x',y')|^2 dx' dy' \right) \left(\int \int |g(x',y')|^2 dx' dy' \right) . \quad (1.1)$$

In Eq.(1.1), $*$ denotes complex conjugation, and the equality holds when $f(x',y')$ is equal to $g(x',y')$. As a result, if the cross-correlation of two functions $f(x',y')$ and $g(x',y')$, which denoted by the left hand side of Eq.(1.1), is normalized by the quantity on the right hand side of Eq.(1.1), one is guaranteed to observe a maximum when

$$f(x',y') = \alpha g(x',y') , \quad (1.2)$$

where α is a constant.

Template matching has been used in applications ranging from automatic character recognition⁵² to scene matching in satellite images⁵⁷ (a complete list of references is given in Chapter 6). Template matching was first implemented digitally, using both Fast Fourier Transforms (FFT's)⁹² and direct computation. In 1964, Vander Lugt¹⁰ demonstrated the first optical implementation of the template matching technique, using a holographic frequency plane filter. Initially, frequency-plane filters were constructed using holographic techniques, using standard recording media such as

photographic plates. More recently, spatial light modulators (for a review see Ref. 94) have been employed in the frequency plane to provide the capacity for real-time updating of the filters. The optical implementation of the frequency plane filter has the advantage of producing the cross-correlation almost instantaneously. The drawbacks of the optical implementation may include the restriction that the input illumination be coherent, or that an incoherent to coherent converter be employed. In addition, the commercially available spatial light modulators have limitations on their resolution, and the time in which they can be updated. Recently, phase-only filters^{49,51} and binary phase-only filters⁵⁰ have been shown to provide good recognition capabilities in certain applications. The holographic phase-only filter has the advantage of providing almost 100% diffraction efficiency, while the binary phase-only filter is more easily implemented using spatial light modulators. In addition, much work has been performed regarding the performance of optical correlations using white light^{12, 19, 20-22}, both for matched filtering²², and rotation-invariant image recognition¹².

A final limitation of the optical implementation is that the correlation output of the filter is not normalized by the quantities indicated by the Schwarz inequality. As a result, it is not guaranteed that one will observe a maximum when the input image and reference function are equal. It must be explicitly stated, however, that it has been well established experimentally that one does indeed observe a maximum in many cases when the input and reference functions are equal.

In practice, the matched filter has been shown to be very sensitive to minute changes in the input image. It has been reported¹⁶ that when a reference object was rotated by 3.5 degrees, or changed scale by 2% with respect to the input, the output signal to noise ratio (SNR) of the correlation peak dropped from 30 dB to 3 dB. (The rate of decrease in SNR increases with the space bandwidth product of the image). As

a result, much work has been directed at obtaining correlation filters that will allow objects to be recognized independent of geometric distortions such as rotation^{12, 17, 18, 34-40, 73}, and scale^{14-16, 24-33, 95}. Additionally, a large body of work exists pertaining to the development of filters that allow objects to be recognized independent of intrinsic variations^{42-48, 51, 74, 96, 97}.

1.2.2 Geometrical-Distortion Invariant Filters

As mentioned before in Section 1.2.1, the approach to the design of distortion-invariant correlation filters has taken essentially two paths. The first is the construction of correlation filters whose correlation output remain unchanged when an input object undergoes changes in rotation, scale, or position. In addition, the filter should have the property that the correlation output should attain a maximum when the reference object is input. These types of filters are usually referred to as invariant filters, and are used to recognize one particular object at a time. The primary filters that exhibit rotation invariance are filters based on the circular-harmonic expansion^{12, 17, 18, 34-40, 73}. These are described in detail in Chapter 4. The most successful filters for scale invariance are based on the Mellin transform¹⁶, and filters or methods for extracting invariant features^{14, 15, 24-33, 95}; these are described in Chapter 5. The most common method of obtaining both rotation and scale invariance is the use of banks of filters. The filter banks typically consist of multiple rotation-invariant filters that are computed for many different scales of the reference object, or vice-versa.

1.2.3 Classification Filters

The second approach to filter synthesis involves the design of correlation filters that tolerate intrinsic distortions, such as out-of-plane rotations, intra-class variations, object occlusions, changing illumination, etc. These filters are sometimes referred to as

classification filters, because their purpose is to designate a particular input object as the member of a particular class. In some cases, researchers^{44,48} have attempted to design filters that can tolerate both intrinsic and geometrical distortions, with some limited success. The most common method of obtaining both rotation-and scale invariance has been to include many rotated, and scaled versions of the reference objects in the training sets. When this technique is used for scale changes greater than 10%, only limited success has been reported⁴⁸.

One of the first attempts at the design of a classification filter is the average filter⁹⁶. The average filter is constructed by averaging together a set of images that consist of the various objects within a certain class. More sophisticated methods include the use of synthetic discriminant functions (SDF's)^{43,44,46,48,51}, composite filters^{42,43,47}, maximum-likelihood filters⁷⁴, and convex-hull filters⁹⁷. In the first three methods mentioned, the correlation filters are designed to produce a specific correlation output when each training-set object is input, with the hope that other objects in the same class, but not in the training set, will produce a similar correlation output. The maximum-likelihood technique⁷⁴ uses statistical decision theory to construct a filter for classifying images. Convex hull filters are constructed from training set data, and provide optimum separability for the closest members of two particular classes.

1.3 Overview of the Thesis

Each chapter of this Thesis details research involving image recognition using a position-sensitive, photon-counting detection system to implement correlation-based methods for image recognition. In Chapter 2, a theoretical formalism is developed that describes the statistical properties of a cross-correlation signal realized by cross correlating a quantum-limited input scene with a complex reference function that is stored in computer memory. This description is performed for both a random, as well as a fixed number of detected photoevents. Hence, one can predict the behavior of the quantum-limited correlation signal when any reference function is utilized. The original work in Chapter 2 involves the extension of the work performed by Morris⁷¹, who described the behavior of a quantum-limited correlation signal realized with a random number of detected photoevents, using a real filter function.

In Chapter 3, the case is examined where the photon-limited input scene is cross-correlated with a classical-intensity image of a reference object. In this case, the correlation output corresponds to that of a matched filter. The number of detected photoevents required to discriminate among a set of detailed objects is predicted using the theory described in Chapter 2, and verified experimentally. In addition, the effect of additive noise on the recognition performance is examined, and a method for reducing the effect of additive noise is described.

In Chapter 4, rotation-invariant image recognition using a circular-harmonic filter is considered. Theoretical predictions are given for the number of detected photoevents required to identify a reference object, and determine its orientation without ambiguity; the theoretical predictions are verified experimentally. In addition, rotation-invariant image recognition in a cluttered environment using circular-harmonic filters is considered. This is performed successfully using a new method for normalizing the

output correlation signal, in near-real time. This is the first description of a method that addresses the problem of properly normalizing the output from a circular-harmonic filter, which has demonstrated inconsistent performance in recognizing objects from within a cluttered environment.

In Chapter 5, the real-time estimation of radial moments of circular-harmonic functions using photon-counting techniques is discussed. This allows many segmented objects to be identified at any orientation or position, and can tolerate changes in scale up to about a factor of two. This is the first real-time method for estimating moment-invariants with sufficient accuracy for pattern recognition applications. Theoretical predictions are made for the number of detected photoevents required to estimate a moment invariant to within a given error; these predictions are verified experimentally.

In Chapter 6, the recognition of objects within cluttered scenes is examined using a two-stage template matching method. The two-stage template matching method is implemented using photon-counting techniques. Computer simulations that demonstrate the recognition performance of the photon-counting implementation of this method for image recognition are performed for several applications. These applications include: the recognition of segmented images, template matching for registration of satellite images, and automatic target recognition in a cluttered environment. The latter is performed for both matched filters, and two-stage filtering using circular-harmonic filters.

Much of the original work presented in this Thesis has been published, or submitted for publication. The work in Chapters 2 and 4 also appears in Ref. 73. Some of the work in Chapter 3 appears in Ref. 76. References 75 and 76 detail some of the early work presented in Chapter 5; a complete description of the work in Chapter

5 has also been submitted for publication (Ref. 78). Finally, much of the work in Chapter 6 appears in Ref. 77.

References for Chapter 1

1. D. H. Ballard and C. M. Brown, *Computer Vision* (Academic, New York, 1984).
2. L. Kanal and A. Rosenfeld, eds. *Progress in Pattern Recognition*, (North-Holland, New York, 1985).
3. IEEE Transactions on Pattern Analysis and Machine Intelligence, published by..
4. Pattern Recognition, published by...
5. This was the number of articles listed when a computer search was done for the key word 'pattern recognition' in the INSPEC data base in October, 1989.
6. B. Kelly and R. Nemhauser, 'Techniques for using the position sensitivity of silicon photodetectors to provide remote machine control,' Proc. IEEE Mach. Tool Conf, 42-57 (1973).
7. T. C. Strand, 'Optical three-dimensional sensing for machine vision,' Opt. Eng. 24, 33-40 (1985).
8. R. C. Gonzalez, *Introduction to Syntactic Pattern Recognition*, (Addison Wesley, Reading, MA 1978).
9. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, (Wiley, New York, 1972).
10. A. B. Van der Lugt, 'Signal detection by complex spatial filtering,' IEEE Trans. IT-10, 139-145 (1964).
11. Y. Hsu and H. H. Arsenault, 'Optical pattern recognition using circular-harmonic expansion,' Appl. Opt. 21, 4016-4019 (1982).
12. K. Mersereau and G. M. Morris, 'Scale, rotation and shift-invariant image recognition,' Appl. Opt. 21, 2338-2342 (1986).
13. W. Steier and R. K. Hoori, 'Optical Hough transform,' Appl. Opt. 25, 2734-2738 (1986).
14. R.A. Athale and J. A. Blodgett, 'Coherent optical moment generator with improved performance,' J. Opt. Soc. Am. 72, 1779-1783 (1982).
15. J. Duvernoy and Y. Sheng, 'Effective optical processor for computing image moments at TV rate: use in handwriting recognition,' Appl. Opt. 26, 2320-2327 (1987).
16. D. Casasent and D. Psaltis, 'Position, rotation and scale invariant optical correlation,' Appl. Opt. 15, 1795-1799 (1976).

17. S. H. Lee, 'Optical implementations of digital algorithms for pattern recognition,' *Opt. Eng.* **25**, 69-75 (1986).
18. G. F. Schils and D. W. Sweeny, 'Rotationally-invariant correlation filtering for multiple objects,' *J. Opt. Soc. Am. A* **3**, 902-909 (1986).
19. S. Wang and N. George, 'Fresnel zone transformations in spatially incoherent illumination,' *Appl. Opt.* **24**, 842-850 (1985).
20. S. Wang and N. George, 'Cosinusoidal transforms in white light,' *Appl. Opt.* **23**, 787-797 (1984).
21. S. Wang and N. George, 'Sine-cosine cascade correlator with real-valued filters,' *Opt. Lett.* **12**, 388-391 (1987).
22. G. M. Morris and N. George, 'Space and wavelength dependence of a dispersion compensated matched filter,' *Appl. Opt.* **18**, 4025-4032 (1985).
23. D. Casasent and A. Mahalamobis, 'Optical iconic filters for large class recognition,' *Appl. Opt.* **26**, 2266-2272 (1987).
24. M.K. Hu, 'Visual pattern recognition by moment invariants,' *IRE Trans. Inf. Theory IT-8*, 179-187 (1961).
25. S. Maitra, 'Moment invariants,' *Proc. IEEE* **67**, 696-699 (1979).
26. S.A. Dudani, K.J. Breeding and R.B. McGee, 'Aircraft identification by moment invariants,' *IEEE Trans. Comput.* **C-26**, 39-45 (1977).
27. R.Y. Wong and E.L. Hall, 'Scene matching with invariant moments,' *Comput. Graphics Image Process.* **8**, 16-21 (1978).
28. R. C. Gonzalez and P. Wintz, *Digital Image Processing*, (Addison-Wesley, Reading, Mass. 1977).
29. D. Casasent and D. Pauly, 'Infrared ship classification using a new pattern recognition concept,' *Proc. Soc. Photo-Opt. Instrum. Eng.* **302**, 126-133 (1981).
30. Y. Abu-Mostafa and D. Psaltis, 'Recognitive aspects of moment invariants,' *IEEE Trans. Patt. Anal. Mach. Intel. PAMI-6*, 698-706 (1984).
31. A. Gostasby, 'Template matching in rotated images,' *IEEE Trans. Patt. Anal. Mach. Intel. PAMI-7*, 338-343 (1985).
32. C. H. Teh and R. T. Chin, 'On digital calculations of moment invariants,' *IEEE Trans. Patt. Anal. Mach. Intel. PAMI-7*, 640-641 (1985).
33. S. S. Reddi, 'Radial and angular moment invariants for image identification,'

- IEEE Trans. Pattern Anal. Machine Intel. PAMI-3, 240-247 (1981).
34. Y.N. Hsu and H.H. Arsenault, 'Rotation-invariant pattern recognition using circular harmonic expansion,' Appl. Opt. 22, 130-132 (1983).
 35. Yunlong Sheng and H.H. Arsenault, 'Method for determining expansion centers and predicting sidelobe levels for circular-harmonic filter,' J. Opt. Soc. Am. A 4, 1793-1797 (1987).
 36. A. Metioui, H.H. Arsenault and L. Leclerc, 'Methods for reducing sidelobes associated with composite filters,' Opt. Commun. 71, 332-336 (1989).
 37. H.H. Arsenault and Yunlong Sheng, 'Properties of the circular-harmonic expansion for rotation-invariant pattern recognition,' Appl. Opt. 25, 3225-3229 (1986).
 38. H.H. Arsenault and Yunlong Sheng, 'Modified composite filter for pattern recognition in the presence of noise with a non-zero mean,' Opt. Commun. 63, 15-20 (1987).
 39. L. Leclerc, Yunlong Sheng and H.H. Arsenault, 'Rotation-invariant phase-only and binary-phase-only correlation,' Appl. Opt. 28, 1251-1256 (1989).
 40. M.T. Manry, 'Compound filter using circular harmonic expansion,' Appl. Opt. 26, 3622-3627 (1987).
 41. R. Wu and H. Stark, 'Rotation-invariant pattern recognition using optimum feature extraction,' Appl. Opt. 24, 179-184 (1985).
 42. H. J. Caufield, 'Linear combinations of filters for character recognition: a unified treatment,' Appl. Opt. 19, 3877-3878 (1980).
 43. B.V.K. Kumar, 'Efficient approach to designing linear combination filters,' Appl. Opt. 22, 1445-1448 (1983).
 44. D. Casasent, 'Unified synthetic discriminant function,' Appl. Opt. 23, 1620-1627 (1984).
 45. B.V.K. Kumar and E. Pochapsky, 'Signal-to-noise ratio considerations in modified matched spatial filters,' J. Opt. Soc. Am. A 3, 777-786 (1986).
 46. B.V.K. Kumar, 'Minimum-variance synthetic discriminant function,' J. Opt. Soc. Am. A 3, 1579-1584 (1986).
 47. R. J. Marks II, J.A. Ritcey, L. E. Atlas and K. F. Cheung, 'Composite matched filter output partitioning,' Appl. Opt. 26, 2274-2278 (1987).
 48. D. Casasent, 'Correlation synthetic discriminant functions,' Appl. Opt. 25, 2343-2350 (1986).
 49. J. L. Horner and P. D. Gianino, 'Phase-only matched filtering,'

- Appl. Opt. 24, 2889-2893 (1985).
50. L. Horner and J. Leger, 'Pattern recognition with binary phase-only filters,' Appl. Opt. 23, 812-816 (1984).
 51. J. L. Horner and P. D. Gianino, 'Applying the phase-only filter to the synthetic discriminant function correlation filter,' Appl. Opt. 24, 851-855 (1985).
 52. J. S. Bomba, 'Alpha-numeric character recognition using local operations,' Fall Joint Comput. Conf., 218-224 (1959).
 53. D. Barnea and H. Silverman, 'A class of algorithms for fast digital image registration,' IEEE Trans. Comput. C-21, 179-186 (1972).
 54. A. Rosenfeld and G. Vanderburg, 'coarse-fine template matching,' IEEE Trans. Sys. Man, Cyber., 104-107 (1977).
 55. G. Vanderburg and A. Rosenfeld, 'Two-stage template matching,' IEEE Trans. Comput. C-26, 384-393 (1977).
 56. A. Goshtasby, S. Gage, and J. Bartholic, 'A two-stage cross correlation approach to template matching,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-6 374-378 (1984).
 57. C. J. Oliver, 'An analysis of template matching in image registration,' Opt. Acta 31, 233-348 (1984).
 58. X. Li and R. C. Dubes, 'The first stage in two-stage template matching,' IEEE Trans. Patt. Anal. Mach. Intell. PAMI-7, 700-707 (1985).
 59. J.L. Turney, T.N. Mudge and R.A. Volz, 'recognizing partially occluded parts,' IEEE Trans. Patt. Anal. Mach. Intell. PAMI-7, 410-421 (1985).
 60. Z. X. Fang, L. Xiao, and L.M. Ni, 'Parallel algorithms for image template matching on hypercube SIMD computers,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-9, 835-841 (1987).
 61. C.I. Soo, J.T. Walker, C.F. Chong, D.H. Damerson and R.F. Pease, 'Template set approach to defect detection and classification for VLSI patterns,' OSA technical Digest Series 12, 18-24 (1987).
 62. D. Casasent and R. Krishnapuram, 'Hough space transformations for discrimination and distortion estimation,' Comput. Vis., Graph. Imag. Proc. 38, 299-316 (1987).
 63. S. Ranka and S. Sahni, 'Image template matching in MIMD hypercube multicomputers,' Proc. 1988 Int. Conf. Parall. Process., Penn. State Univ., 15-19 (1988).
 64. D. Casasent and D. Psaltis, 'Hybrid processor to compute invariant moments for pattern recognition,' Opt. Lett. 5, 395-398 (1980).

65. K. Wagner and D. Psaltis, 'Real-time computation of moments with acousto-optics,' *Proc. SPIE* 352, 82-89 (1982).
66. B.V.K. Kumar and C.A. Rahencamp, 'Calculation of geometric moments from Fourier plane intensities,' *Appl. Opt.* 6, 997-1007 (1986).
67. H. L. Kasdan and N. George, 'Linewidth measurement by diffraction pattern sampling,' *Proc IEEE* 58, 198-216 (1976).
68. D. Venable, *Pattern classification using diffraction-pattern sampling* Ph.D. Thesis, University of Rochester, (1989).
69. New focal plane architecture and transform for fast target recognition,' *Appl. Opt.* 28, 3810-3813 (1989).
70. G.M. Morris, 'Optical computing using Monte Carlo methods,' *Opt. Eng.* 24, 86-90 (1985).
71. G. M. Morris, 'Scene matching at low light levels: a computer simulation,' *Appl. Opt.* 23, 3152-3159 (1984).
72. G. M. Morris, M. N. Wernick and T. A. Isberg, 'Image correlation at low light levels,' *Opt. Lett.* 10, 315-317 (1985).
73. T.A. Isberg and G.M. Morris, 'Rotation-invariant image recognition at low light levels,' *J. Opt. Soc. Am. A* 3, 954-963 (1986).
74. M.N. Wernick and G.M. Morris, 'Image classification at low light levels,' *J. Opt. Soc. A* 3, 2179-2187 (1986).
75. T. A. Isberg and G. M. Morris, 'Monte-Carlo estimation of moment invariants for pattern recognition,' *Proc. SPIE* 938 365-371 (1988).
76. G. M. Morris, T. A. Isberg and M. N. Wernick, 'Pattern recognition using photon-limited images,' *Proc. SPIE* 960, 86-102 (1989).
77. T. A. Isberg and G. M. Morris, 'Two-Stage Template Matching using quantum-limited images,' *Proc. SPIE* 976, 160-167 (1988).
78. T. A. Isberg and G. M. Morris, 'Real-time estimation of moment invariants for pattern recognition,' submitted to *J. Opt. Soc. Am. A*, November 1989.
78. T. H. Tanaka, D. H. Ballard, S. Tsuji, and M. Curtiss, 'Parallel polyhedral shape recognition,' *Proc. CVPR: IEEE Comp. Soc. Conf. Comput. Vis. and Patt. Recog.* 19-23 (1985).
79. D. H. Ballard and C. M. Brown, 'Vision,' *BYTE* 10, 245-261 (1985).
80. T. J. Olson, L. Bukys, and C. M. Brown, 'Low level image analysis on an MIMD architecture,' *Proc. First Int. Conf. Comput. Vis.* 8-11 (1987).

81. R. G. Buser and F. Warren, eds., *Infrared Sensors and Sensor Fusion*, Proc. SPIE 782, (1987).
82. S. Rogers, C. Tong, M. Kabrisky and J. Mills, 'Multisensor fusion of ladar and passive imagery for target segmentation,' *Opt. Eng.* 28, 881-886 (1989).
83. N. Nandakumar and J. Aggarwal, 'The artificial intelligence approach to pattern recognition- a perspective and overview,' *Patt. Recog.* 18, 383-389 (1985).
84. D. E. Rumelhart, J. L. McClelland, *et. al.*, *Parallel Distributed Processing*, vol. 1, chapter 8, (MIT Press, Cambridge, 1986).
85. E. Paek and D. Psaltis, 'Optical associative memory using Fourier transform holograms,' *Opt. Eng.* 26, 428-433 (1987).
86. L. Saaf and G. M. Morris, 'Back-propagation filters for image classification,' submitted to *Appl. Opt.*, Nov. (1989).
87. ITT Electro-Optical Products Division, Tube and Sensors Laboratories, Fort Wayne, IN. Detector model F4146M.
88. Surface Science Laboratories Inc., Mountain View, CA. Model 2401 position computer.
89. L. Mandel, E. C. G. Sudarshan and E. Wolf, 'Theory of photoelectric detection of light fluctuations,' *Proc. Phys. Soc.* 84, 435-444 (1964).
90. M. H. Kalosh and P. A. Whitlock, *Monte Carlo Methods, Vol. 1* (Wiley, New York, 1986).
91. G. Arfken, *Mathematical Methods for Physicists*, (Academic, New York, 1970) p.445.
92. E. Brigham, *The Fast Fourier Transform*, (Prentis-Hall, Englewood Cliffs, 1974).
93. For example, Eastman Kodak 131 and 649 photographic plates.
94. T. Hudson, 'The current technological status of spatial light modulators,' U.S. Army Missile Command Technical Report RD-RE-88-5 (1988).
95. Y. Sheng, 'Fourier-Mellin spatial filters for invariant pattern recognition,' *Opt. Eng.* 28, 494-500 (1989).
96. C. Hester and D. Casasent, 'Optical pattern recognition using average filters to produce discriminant hypersurfaces,' *Proc. SPIE* 201, 77082 (1979).
97. M. Wernick, Ph.D. Thesis, Univ. of Rochester, (1989) Chap. 3.

Chapter 2.

Correlation with a Deterministic Reference Function

2.1 Introduction

The approach to pattern recognition using photon-limited images can be described in terms of a linear filtering formalism. In this chapter, a general treatment is presented for the statistical properties of a correlation signal realized by cross correlating a photon-limited input scene with a reference function that is stored in computer memory. Morris¹ performed the analysis for the case when the reference function was real, and the number of detected photoevents was Poisson distributed. The theoretical predictions were verified using complicated objects in a computer simulation². For the work in this thesis, it was necessary to consider the most general case, where the reference function is taken to be complex, and the number of detected photoevents can be either Poisson distributed or fixed. In Section 2.2, some statistics describing the detection of photon-limited images are briefly reviewed. In Section 2.3, the statistical behavior of the correlation signal is described for the case when the reference function is complex. Approximate expressions are given for the probability density functions of the correlation signal when the number of detected photoevents, N , is Poisson-distributed, as well as for the case when N is fixed. In Section 2.4, the results are specialized to the case when the reference function is real (these results agree with those obtained by Morris¹). The theoretical predictions for the probability density functions and statistical moments are used to determine the accuracy of the image recognition process for a given input and reference. The formalism for this detection process is reviewed in Section 2.5. Finally, the effect of additive noise on the statistics

of the correlation signal is treated in Section 2.6. The theoretical predictions for the statistical behavior of the photon-limited correlation signal that are presented in this chapter are experimentally verified in later chapters.

2.2 Detection of Photon-Limited Images

A low-light level input image $f^\dagger(x',y')$ can be represented as a collection of Dirac delta functions,

$$f^\dagger(x',y') = \sum_{i=1}^N \delta(x' - x_i, y' - y_i) \quad , \quad (2.1)$$

where (x_i, y_i) represent the coordinate location of the i^{th} detected photoevent, and N is the total number of detected photoevents in the reference scene window. In this subsection, we briefly review some of the photoevent detection statistics required for the description of the statistical behavior of the photon-limited correlation signal.

From the theory of photodetection³, the probability of emission of an electron in a small time interval Δt from a photosurface of differential area ΔA is found to be

$$P(x, y, t) \Delta t \Delta A = \frac{\eta f(x, y) \Delta t \Delta A}{h \bar{\nu}} \quad , \quad (2.2)$$

where η denotes the quantum efficiency of the detector, $f(x, y)$ is a classical measure of the instantaneous image intensity, h is Planck's constant, $\bar{\nu}$ denotes the mean frequency of the quasi-monochromatic incident illumination and (x, y) denotes the position of the area ΔA .

Given Eq. (2.2) it follows from first principles³ that the probability distribution of emission of N photoelectrons in a finite time interval $[t, t+\tau]$ from a detector of area A is a Poisson process,

$$P(N) = \frac{\left[\int_t^{t+\tau} \int_A dx dy \lambda(x, y; t') \right]^N}{N!} \exp \left\{ - \int_t^{t+\tau} \int_A dx dy \lambda(x, y; t') \right\}, \quad (2.3)$$

in which the rate function $\lambda(x, y; t)$ is

$$\lambda(x, y; t) = \frac{\eta f(x, y; t)}{h\nu} \quad (2.4)$$

Generally, $f(x, y; t)$ is a random process, and one must take an ensemble over the distribution in Eq. (2.3) to predict the observed counting distribution.

If $f(x, y; t)$ does not fluctuate significantly, as is the case of the illumination provided by a well stabilized, single mode laser, or for polarized, quasi-monochromatic thermal radiation where τ is much larger than the coherence time of the source, the counting distribution becomes Poisson distributed^{3,4},

$$P(N) = \frac{\bar{N}^N}{N!} e^{-\bar{N}} \quad (2.5)$$

where

$$\bar{N} = \frac{[\eta\tau/h\nu]}{\int \int_A dx dy f(x, y)} \quad (2.6)$$

Using Eqs. (2.2) and (2.6), the probability of the photoevent coordinates can be written as

$$p(x, y) = \frac{f(x, y)}{\int \int_A f(x, y) dx dy} \quad (2.7)$$

2.3 Correlation with a Complex-Valued Reference Function

2.3.1 Poisson-Distributed Number of Detected Photoevents

The cross-correlation signal

$$C(x, y) = \int \int_A f^\dagger(x', y') R(x + x', y + y') dx' dy' \quad (2.8)$$

between a reference image $R(x, y)$ stored in computer memory and the photon-limited input image $f^\dagger(x', y')$ is given by (using Eqs. (2.1) and (2.8))

$$C(x, y) = \sum_{i=1}^N R(x + x_i, y + y_i) \quad (2.9)$$

This process is shown schematically in Fig. 2.1. Hence, the photon-limited correlation signal is a random function, since the photoevent coordinates are independent random variables. As mentioned earlier, the number of detected photoevents N may also be random, depending upon how the experiment is performed.

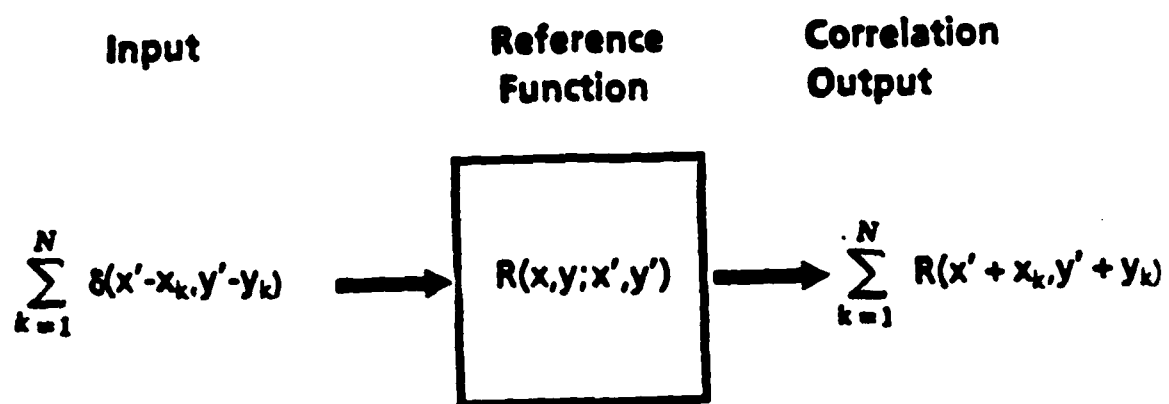


Figure 2.1 System diagram for quantum-limited image recognition

One must be able to theoretically predict the statistical behavior of the photon-limited correlation signal if this photon-counting method for image recognition is to be implemented. In particular, it is necessary to obtain expressions for the first and second moments of the correlation signal, as well as an expression for the probability density function. The characteristic function of a random variable often provides an elegant means to obtain closed form solutions to many problems in probability theory, particularly those associated with sums of random variables. That is the approach employed here. For convenience, we define the complex photon-limited correlation signal $C(x,y)$ as

$$C(x,y) = C'(x,y) + i C''(x,y) \quad , \quad (2.10)$$

where $C'(x,y)$ and $C''(x,y)$ denote the real and imaginary parts of the correlation signal respectively. The joint characteristic function $\Phi(\omega',\omega'')$ is defined as⁵

$$\Phi(\omega', \omega'') = \langle \exp [i [(\omega' C'(x,y) + \omega'' C''(x,y))]] \rangle \quad , \quad (2.11)$$

where $\langle \dots \rangle$ denotes an ensemble average. When the counting distribution in Eq. (2.5) is valid, the characteristic function can be derived in a manner similar to that for shot noise⁶.

To derive the characteristic function, we start by dividing the reference scene $R(x,y)$ into small differential areas ΔA_{ij} , where ΔA_{ij} is defined as

$$\Delta A_{ij} \equiv [(x_i, y_j), (x_i + \Delta x, y_j + \Delta y)] \quad . \quad (2.12)$$

The lengths Δx and Δy are sufficiently small enough to insure that the value of the reference function will be constant over the area. Hence, in the intervals defined by

$$[x_i \leq x \leq x_i + \Delta x, y_j \leq y \leq y_j + \Delta y] \quad , \quad (2.13)$$

the value of the reference function is approximately constant;

$$R(x, y) \approx R(x + x_i, y + y_j) \quad (2.14)$$

The contribution to $C'(x, y)$ and $C''(x, y)$ from each area $\Delta A_{i,j}$ is then

$$\Delta C'_{i,j}(x, y) = \text{Re} \{ R(x + x_i, y + y_j) \} \cdot \Delta N_{i,j} \quad (2.15)$$

and

$$\Delta C''_{i,j}(x, y) = \text{Im} \{ R(x + x_i, y + y_j) \} \cdot \Delta N_{i,j} \quad (2.16)$$

respectively. In Eqs. (2.15) and (2.16), $\text{Re}\{.. \}$ and $\text{Im}\{.. \}$ denote the real and imaginary parts of the reference function, and $\Delta N_{i,j}$ denotes the number of detected photoevents in the area $\Delta A_{i,j}$ in the time interval τ . We assume that the areas $\Delta A_{i,j}$ do not overlap, and are small enough such that the number of detected photoevents in each differential area are statistically independent. Using Eqs. (2.15) and (2.16), $C'(x, y)$ and $C''(x, y)$ can be written as

$$C'(x, y) = \sum_{i,j} \Delta C'_{i,j}(x, y) \quad (2.17)$$

and

$$C''(x, y) = \sum_{i,j} \Delta C''_{i,j}(x, y) \quad (2.18)$$

respectively. Substituting Eqs. (2.15)-(2.18) into Eq. (2.11), the joint characteristic function of $C'(x, y)$ and $C''(x, y)$ is written as

$$\Phi(\omega', \omega'') = \langle \exp [i \sum_i \sum_j (\omega' \operatorname{Re} \{R(x + x_i, y + y_j)\} + \omega'' \operatorname{Im} \{R(x + x_i, y + y_j)\}) \Delta N_{i,j}] \rangle . \quad (2.19)$$

If the number of detected photoevents $\Delta N_{i,j}$ is Poisson distributed (see Eq. (2.5)) with rate parameter

$$\Delta \bar{N}_{i,j} = \frac{\eta \tau}{h\nu} f(x_i, y_j) \Delta x \Delta y , \quad (2.20)$$

the characteristic function becomes

$$\Phi(\omega', \omega'') = \sum_{\Delta N_{i,j}=0}^{\infty} \frac{\bar{N}_{i,j}^{\Delta N_{i,j}} e^{-\bar{N}_{i,j}}}{(\Delta N_{i,j})!} \exp [i \sum_i \sum_j (\omega' \operatorname{Re} \{R(x + x_i, y + y_j)\} + \omega'' \operatorname{Im} \{R(x + x_i, y + y_j)\}) \Delta N_{i,j}] . \quad (2.21)$$

Re-grouping terms in Eq. (2.21), and using the infinite series definition of $\exp\{x\}$, Eq. (2.21) is re-written as

$$\Phi(\omega', \omega'') = e^{-\bar{N}} \exp \{ \bar{N}_{i,j} \exp [i \sum_i \sum_j (\omega' \operatorname{Re} \{R(x + x_i, y + y_j)\} + \omega'' \operatorname{Im} \{R(x + x_i, y + y_j)\}) \Delta N_{i,j}] \} . \quad (2.22)$$

Substituting Eq. (2.20) into Eq. (2.22), the expression for $\Phi(\omega', \omega'')$ becomes

$$\Phi(\omega', \omega'') = \exp \left\{ \frac{\eta \tau}{h\nu} \sum_i \sum_j f(x_i, y_j) \Delta x \Delta y [\exp i (\omega' \operatorname{Re} \{R(x + x_i, y + y_j)\} + \omega'' \operatorname{Im} \{R(x + x_i, y + y_j)\}) - 1] \right\} . \quad (2.23)$$

In the limit as Δx and Δy approach zero, the sums tend to integrals. In addition, if one multiplies and divides the argument of the exponential by

$$\iint_A f(x', y') dx dy, \quad (2.24)$$

the expression for $\Phi(\omega', \omega'')$ becomes

$$\begin{aligned} \Phi(\omega', \omega'') = \exp \left\{ \bar{N} \iint_A p(x', y') [\exp i(\omega' \text{Re}\{R(x + x', y + y')\}) \right. \\ \left. + \omega'' \text{Im}\{R(x + x', y + y')\}) - 1] \right\} \end{aligned} \quad (2.25)$$

where $p(x', y')$ is given Eq (2.7).

The joint moments of the correlation signal are computed from the characteristic function using the relation

$$\langle C'^m(x, y) C''^n(x, y) \rangle = \frac{(-i)^{m+n} \partial^{m+n} \Phi(\omega', \omega'')}{\partial \omega'^m \partial \omega''^n} \bigg|_{\omega' = \omega'' = 0} \quad (2.26)$$

Using Eqs. (2.25) and (2.26), the mean values of the real and imaginary parts of the correlation signal, $\langle C'(x, y) \rangle$ and $\langle C''(x, y) \rangle$, are found to be⁷

$$\langle C'(x, y) \rangle = \bar{N} \iint_A p(x', y') \text{Re}\{R(x + x', y + y')\} dx' dy' \quad (2.27)$$

and

$$\langle C''(x, y) \rangle = \bar{N} \iint_A p(x', y') \text{Im}\{R(x + x', y + y')\} dx' dy' \quad (2.28)$$

respectively. The variances, σ'^2 and σ''^2 are given by

$$\sigma'^2 = \bar{N} \iint_A p(x', y') [\text{Re}\{R(x + x', y + y')\}]^2 dx' dy' \quad (2.29)$$

and

$$\sigma''^2 = \bar{N} \int \int_A p(x', y') [\text{Im}\{R(x + x', y + y')\}]^2 dx' dy' \quad (2.30)$$

respectively. The correlation coefficient ρ , defined as

$$\rho = \frac{\langle C C'' \rangle - \langle C' \rangle \langle C'' \rangle}{\sigma' \sigma''} \quad (2.31)$$

is found to be

$$\rho = \frac{\bar{N} \int \int_A p(x', y') \text{Im}\{R(x + x', y + y')\} \text{Re}\{R(x + x', y + y')\} dx' dy'}{\sigma' \sigma''} \quad (2.32)$$

The correlation between the real and imaginary parts of the correlation signal occurs because the real and imaginary parts of the reference function may be correlated; the correlation is not introduced because of any correlation in the detection of the photoevents. In certain applications, the correlation coefficient ρ has been observed to be quite small, and can be neglected in the calculation of the joint probability density function [shown later in Appendix A, Eq. (A15)].

It is important to note that the mean value of both the real and imaginary parts of the photon-limited correlation signal, denoted by $\langle C' \rangle$ and $\langle C'' \rangle$ in Eqs. (2.27) and (2.28) respectively, are directly proportional to the cross-correlation of the classical intensity input scene $f(x', y')$ with the reference function $R(x', y')$. This fact enables one to estimate the high-light-level cross-correlation signal using a small number of detected photoevents. The variance in the correlation signal is due to the statistical fluctuations inherent in the low-light-level input scene.

One must now compute the joint probability density function for the real and imaginary parts of the correlation signal. The joint probability density function for the

real and imaginary parts of the correlation signal is obtained by taking the inverse Fourier transform of the characteristic function. Unfortunately, the Fourier inverse of Eq. (2.25) is very difficult to obtain in general⁸. If the average number of detected photoevents \bar{N} is large, it is straightforward to derive a limit form for the characteristic function and hence for the probability density function⁷.

Note that in Eqs. (2.27)-(2.32), the moments of the correlation signal become infinite as \bar{N} becomes large. This difficulty is easily avoided by making the following standard change of variables: Let

$$\zeta(x, y) = \zeta'(x, y) + i \zeta''(x, y), \quad (2.33)$$

where

$$\zeta'(x, y) = \frac{C'(x, y) - \langle C'(x, y) \rangle}{\sigma'} \quad (2.34)$$

and

$$\zeta''(x, y) = \frac{C''(x, y) - \langle C''(x, y) \rangle}{\sigma''} \quad (2.35)$$

respectively. The joint characteristic function of $\zeta'(x, y)$ and $\zeta''(x, y)$ is then given by

$$\begin{aligned} \Phi_{\zeta}(\omega', \omega'') &= \exp \left\{ i \left(\frac{\langle C'(x, y) \rangle}{\sigma'} + \frac{\langle C''(x, y) \rangle}{\sigma''} \right) \right\} \\ &\exp \left\{ \bar{N} \int_A \int p(x', y') \left[\exp i \left(\frac{\omega'}{\sigma'} \operatorname{Re} \{ R(x + x', y + y') \} \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{\omega''}{\sigma''} \operatorname{Im} \{ R(x + x', y + y') \} \right) - 1 \right] \right\} . \end{aligned} \quad (2.36)$$

To derive a large \bar{N} limit for Eq. (2.36), we expand the inner exponential in a power series:

$$\begin{aligned}
& \exp \left[i \left(\frac{\omega'}{\sigma'} \operatorname{Re} \{R(x+x', y+y')\} + \frac{\omega''}{\sigma''} \operatorname{Im} \{R(x+x', y+y')\} \right) \right] - 1 \\
&= 1 - 1 + i \left[\frac{\omega'}{\sigma'} \operatorname{Re} \{R(x+x', y+y')\} + \frac{\omega''}{\sigma''} \operatorname{Im} \{R(x+x', y+y')\} \right] \\
&\quad - \frac{1}{2} \left[\left(\frac{\omega'}{\sigma'} \operatorname{Re} \{R(x+x', y+y')\} \right)^2 + \left(\frac{\omega''}{\sigma''} \operatorname{Im} \{R(x+x', y+y')\} \right)^2 \right. \\
&\quad \left. + \left(\frac{\omega' \omega'' \operatorname{Re} \{R(x+x', y+y')\} \operatorname{Im} \{R(x+x', y+y')\}}{\sigma' \sigma''} \right) \right] \\
&\quad - \frac{i}{6} \left[\frac{\omega'}{\sigma'} \operatorname{Re} \{R(x+x', y+y')\} + \frac{\omega''}{\sigma''} \operatorname{Im} \{R(x+x', y+y')\} \right]^3 \\
&\quad + \dots
\end{aligned} \tag{2.37}$$

If one then integrates the series term by term, and uses Eqs. (2.29) and (2.30), one sees immediately that the terms proportional to $1/\sigma'$ and $1/\sigma''$ cancel with the first exponential in Eq. (2.36). The next three terms become after integration:

$$\begin{aligned}
& \frac{\bar{N}}{2} \int_A dx' dy' p(x', y') \left[\left(\frac{\omega'}{\sigma'} \operatorname{Re} \{R(x+x', y+y')\} \right)^2 \right. \\
& \quad \left. + \left(\frac{\omega''}{\sigma''} \operatorname{Im} \{R(x+x', y+y')\} \right)^2 + \left(\frac{\omega' \omega'' \operatorname{Re} \{R(x+x', y+y')\} \operatorname{Im} \{R(x+x', y+y')\}}{\sigma'} \right) \right] .
\end{aligned} \tag{2.38}$$

Since both σ' and σ'' are proportional to $\bar{N}^{\frac{1}{2}}$, the terms in the series proportional to $(1/\sigma)^3$ and all following terms tend to zero in the limit as \bar{N} approaches infinity. Thus, for large \bar{N} , the characteristic function becomes

$$\Phi_c(\omega', \omega'') = \exp - \left\{ \frac{1}{2} (\omega'^2 + 2\rho\omega'\omega'' + \omega''^2) \right\} . \tag{2.39}$$

Note that Eq. (2.39) is the equation for the joint characteristic function of two random variables with a bi-variate normal probability distribution. Hence, when \bar{N} is large, the real and imaginary parts of the correlation signal are approximately jointly normal distributed, with the form⁷

$$P(C', C'') = \frac{1}{2\pi\sigma'\sigma''(1-\rho^2)^{\frac{1}{2}}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(C' - \langle C' \rangle)^2}{\sigma'^2} - 2\rho \frac{(C' - \langle C' \rangle)(C'' - \langle C'' \rangle)}{\sigma'\sigma''} + \frac{(C'' - \langle C'' \rangle)^2}{\sigma''^2} \right] \right\}, \quad (2.40)$$

where $\langle C' \rangle$, $\langle C'' \rangle$, σ' , σ'' are given in Eqs. (2.27)-(2.30). Equation (2.40) serves as the basis for many of the statistical calculations involving the photon-limited correlation signal, and is referred to quite frequently in later chapters. In particular, it is often necessary to determine the probability density function for random variables that are based on the photon-limited correlation. For example, it is often necessary to determine the probability density function for the modulus of the complex correlation signal (see Chapters 4 and 5). Equation (2.40) serves as the basis in the computation of this density.

2.3.2 N Fixed

In the previous section, it is assumed that the correlation signal is realized by detecting photoevents for a fixed time interval τ . Hence, provided that the assumptions concerning the illuminating radiation are correct, the number of detected photoevents is a Poisson distributed random variable. If the correlation signal is realized by detecting a fixed number of photoevents, one source of the random fluctuations in the value of

the correlation signal is removed. The effect on the statistics of the observed correlation signal is found to be a decrease in the variance of the real and imaginary parts of the correlation signal. This is demonstrated via the following analysis.

If the cross correlation signal of a photon-limited input scene with a reference function is performed using a fixed number of photoevents N (see Eq. (2.9)), then, by the central limit theorem¹⁰, the probability density function of the real and imaginary parts of the correlation will approach a normal distribution, assuming the number of detected photoevents is large, since the photoevent coordinates are statistically independent. The first and second order moments of the correlation signal can be computed directly. The mean value of the real and imaginary parts of the correlation signal are given by

$$\langle C'(x, y) \rangle = \left\langle \sum_{i=1}^N \text{Re} \{ R(x + x_i, y + y_i) \} \right\rangle, \quad (2.41)$$

and

$$\langle C''(x, y) \rangle = \left\langle \sum_{i=1}^N \text{Im} \{ R(x + x_i, y + y_i) \} \right\rangle, \quad (2.42)$$

where $\langle \dots \rangle$ denotes an ensemble average. Using Eq. (2.7) for the probability density function of the photoevent coordinates, and performing the ensemble average term by term, one obtains

$$\langle C'(x, y) \rangle = N \int_A \int p(x', y') \text{Re} \{ R(x + x', y + y') \} dx' dy' \quad , \quad (2.43)$$

and

$$\langle C''(x, y) \rangle = N \int_A \int p(x', y') \text{Im} \{ R(x + x', y + y') \} dx' dy' \quad , \quad (2.44)$$

respectively. Note that the mean value of the correlation signal is the same as in the case when N is Poisson distributed, if one replaces \bar{N} with N . The variance of the real and imaginary parts, denoted by σ'^2 and σ''^2 respectively, is calculated in a similar manner. By definition, σ'^2 is given by

$$\sigma'^2 = \langle C'^2(x, y) \rangle - \langle C'(x, y) \rangle^2 \quad (2.45)$$

Computing $\langle C'^2(x, y) \rangle$ directly, using Eq. (2.7), one finds

$$\begin{aligned} \langle C'^2(x, y) \rangle &= N \int_A \int p(x', y') \text{Re} \{ R(x + x', y + y') \}^2 dx' dy' \\ &+ (N^2 - N) \left[\int_A \int p(x', y') \text{Re} \{ R(x + x', y + y') \} \right]^2 \quad (2.46) \end{aligned}$$

Hence, using Eqs. (2.43), (2.45) and (2.46), one finds

$$\begin{aligned} \sigma'^2 &= N \int_A \int p(x', y') [\text{Re} \{ R(x + x', y + y') \}]^2 dx' dy' \\ &- N \left[\int_A \int p(x', y') R(x + x', y + y') \right]^2 \quad (2.47) \end{aligned}$$

Noting that the second term in Eq. (2.47) is $(1/N) \langle C'(x, y) \rangle^2$, one obtains

$$\sigma'^2 = N \int_A \int p(x', y') [\text{Re} \{ R(x + x', y + y') \}]^2 dx' dy' - \frac{\langle C'(x, y) \rangle^2}{N} \quad (2.48)$$

In a similar fashion, one obtains for σ''^2 ,

$$\sigma'^2 = N \int \int_A p(x', y') [\text{Im}\{R(x + x', y + y')\}]^2 dx' dy' - \frac{\langle C'(x, y) \rangle^2}{N} \quad (2.49)$$

The correlation coefficient ρ is found to be

$$\rho = \left[\frac{N}{\sigma' \sigma''} \right] \int \int_A p(x', y') \text{Re}\{R(x + x', y + y')\} \text{Im}\{R(x + x', y + y')\} dx' dy' \quad (2.50)$$

$$= \frac{\langle C'(x, y) \rangle \langle C''(x, y) \rangle}{\sigma' \sigma''} ,$$

where σ'^2 and σ''^2 are given in Eqs. (2.48) and (2.49). One may also note that the first terms in Eqs. (2.48) and (2.49) are the variance for the case when N is Poisson distributed (if \bar{N} is replaced by N). This demonstrates the reduction in the variance that is observed when the number of detected photoevents is fixed.

2.4. Correlation with a Real Reference Function

2.4.1 N Poisson Distributed

The statistics for the case when a photon-limited input scene is correlated with a reference function that is entirely real were first computed by Morris¹. Here, we obtain the same results directly from the expressions given in the previous section by letting the complex part of the reference function equal zero. The moments of the correlation signal are obtained by letting the imaginary part of the reference function go to zero in Eqs. (2.27)-(2.32). Clearly, only the moments for the real part are non-zero; hence the "prime" notation is eliminated. By inspection, one obtains for the expected value and variance,

$$\langle C(x, y) \rangle = \bar{N} \int \int_A p(x', y') R(x + x', y + y') dx' dy' \quad (2.51)$$

and

$$\sigma^2 = \bar{N} \int \int_A p(x', y') R^2(x + x', y + y') dx' dy' \quad , \quad (2.52)$$

respectively.

If one lets the moments for the complex part of the correlation signal be zero in Eq. (2.40), one finds that the probability density function for the real correlation signal is a monovariate normal distribution, given by

$$P(C(x, y)) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \left\{ \frac{C(x, y) - \langle C(x, y) \rangle}{2\sigma^2} \right\}^2 \quad . \quad (2.53)$$

In Eq. (2.53), $\langle C(x, y) \rangle$ and σ^2 are given by Eqs. (2.51) and (2.52) respectively.

2.4.2 N Fixed

The moments of probability density function for the photon-limited correlation signal realized by cross-correlating a real reference function with photon-limited input scenes containing a fixed number of detected photoevents N can also be obtained by specializing the results from the complex case. In this case, we let the imaginary part of the reference function be zero in Eqs. (2.42)-(2.44), and in Eqs. (2.45)-(2.48). One immediately obtains for the expected value and variance

$$\langle C(x, y) \rangle = N \int \int_A p(x', y') R(x + x', y + y') dx' dy' \quad , \quad (2.54)$$

and

$$\sigma^2 = N \int \int_A p(x', y') R^2(x + x', y + y') dx' dy' - \frac{\langle C \rangle^2}{N}, \quad (2.55)$$

respectively.

The probability density function of the correlation signal for the case when N is fixed and the reference function is real has the same form as Eq. (2.53),

$$P(C(x,y)) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \left\{ \frac{C(x,y) - \langle C(x,y) \rangle}{2\sigma^2} \right\}^2 \quad (2.56)$$

with $\langle C(x,y) \rangle$ and σ^2 given by Eqs. (2.54) and (2.55), respectively.

2.5 Probability of Detection and False Alarm

The exact number of detected photoevents required to discriminate among a set of given images can be determined using the statistical theory of hypothesis testing¹¹. On the basis of the photon-limited cross-correlation $C(x,y)$, one must choose between two hypotheses: the null hypothesis H_0 -- input image $f(x,y)$ is *not* the same as the reference image $R(x,y)$, or the positive hypothesis, H_1 -- the input image $f(x,y)$ is the same as the reference image $R(x,y)$. Under hypothesis H_0 , the probability density function for $C(x,y)$ is denoted by $P_0(C(x,y)) = P[C(x,y) | f(x,y) = N(x,y)]$, where $N(x,y)$ is some noise, or false image. Under hypothesis H_1 , the density function of C is denoted by $P_1(C(x,y)) = P[C(x,y) | f(x,y) = R(x,y)]$, where $R(x,y)$ is the reference image.

As indicated earlier the observer sets a threshold C_T for the correlation signal. If the estimate for $C(x,y) > C_T$, hypothesis H_0 is chosen; if $C(x,y) < C_T$, hypothesis H_1 is chosen. However, because of the statistical nature of the estimate for $C(x,y)$, the observer occasionally makes an error, regardless of the value chosen for C_T . The probability of choosing H_1 when H_0 is true is called the probability of false alarm and is given by

$$P_{fa} = \int_{C_T}^{\infty} P_0(C(x,y)) dC(x,y) \quad (2.57)$$

The probability of choosing H_1 when H_1 is true is called the probability of detection, and is given by

$$P_d = \int_{C_T}^{\infty} P_1(C(x,y)) dC(x,y) \quad (2.58)$$

The overlap areas associated with P_d and P_{fa} are shown in Fig. 2.2

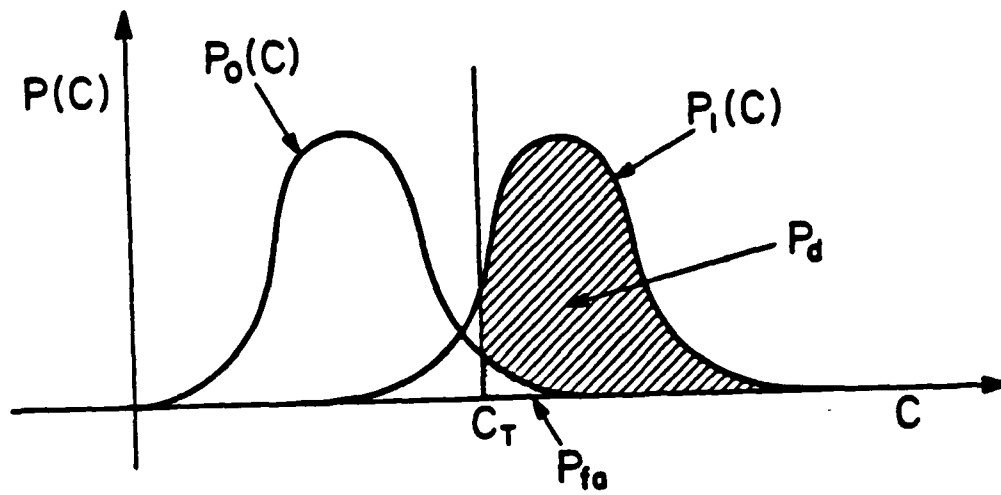


Figure 2.2 Probability density functions under hypotheses H_0 and H_1 .

To determine the number of detected photoevents N that are required to calculate the correlation signal within a given error, one must specify the required probabilities of detection and false alarm for a given application. Next, one calculates the probability density functions for the photon-limited estimates for $C(x,y)$ for the input and reference images for a starting number of detected photoevents N ; the functional form of this density function is given in Section 2.3 or 2.4. One then computes P_d and P_{fa} using the appropriate density functions, and compares these values with the required ones. If the calculated values for P_d and P_{fa} do not meet the desired specifications, one increases the value for N , and repeats the process until the required probabilities of detection and false alarm are achieved.

2.6 Summary of Chapter 2

A means for theoretically predicting the statistical behavior of a correlation signal realized by cross-correlating a photon-limited input scene with a complex reference function stored in computer memory is presented. This is done for the case when the number of detected photoevents, N , used to realize the correlation signal Poisson distributed, as well as for the case when the number of detected photoevents is fixed. The results for the moments of the correlation signal are given in Eqs. (2.27)-(2.32) for the case when N is Poisson distributed; when N is fixed, the statistical moments are given by Eqs. (2.43)-(2.44), and (2.48)-(2.50). The probability density function for the correlation signal in both cases is bi-variate normal, with the form given in Eq. (2.40).

When the reference function is entirely real, and the number of detected photoevents N is Poisson distributed, the expressions for the statistical moments are given by Eqs. (2.51) and (2.52); the probability density function is given in Eq. (2.53). These results are in agreement with those first obtained by Morris¹. When N is fixed,

the mean value and variance of the correlation signal are given by Eqs. (2.54) and (2.55) respectively, with the probability density function shown in Eq. (2.56).

References of Chapter 2

1. G. M. Morris, 'Scene matching at low light levels,' J. Opt. Soc. Am. A **1**, 482-488 (1984).
2. G. M. Morris, 'Scene matching at low light levels: a computer simulation,' Appl. Opt. **23**, 3152-3159 (1984).
3. L. Mandel, E. C. G. Sudarshan and E. Wolf, 'Theory of photoelectric detection of light fluctuations,' Proc. Phys. Soc. **84**, 435-444 (1964).
4. M. Bertolotti, 'Photon Statistics,' in *Photon Correlation and Light Beating Spectroscopy*, H. Z. Cummins and E. L. Pike, eds. (Plenum, New York, 1974), Chap. 2.
5. A. Papoulis, *Probability Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1965), p.213.
6. *Ibid.* p. 382.
7. T.A. Isberg and G.M. Morris, 'Rotation-invariant image recognition at low light levels,' J. Opt. Soc. Am. A **3**, 954-963 (1986).
8. K. S. Miller, *Multivariate Distributions* (Krieger, New York, 1975) p. 28.
9. Reference 5, p. 225.
10. B. V. Gnendenko, *Theory of Probability* (Chelsea, New York, 1963) Chap. 8.
11. A. Sage and J. Melsa, *Estimation Theory with Applications to Communication and Control* (McGraw-Hill, New York, 1971) p. 118-120.

Chapter 3.

Image Correlation at Low Light Levels

3.1 Introduction

The results of the previous chapter are applicable for general reference functions and input scenes. In this chapter, we study the case where the photon-limited input scene is cross correlated with the classical-intensity image of the reference object. In this case, the correlation output $C(x,y)$ in Eq. (2.9) corresponds to that of a matched filter. In this chapter, we begin in Section 3.2 by briefly describing the laboratory system used to acquire the photon-limited images and perform the image correlations. Next, in Section 3.3 the number of detected photoevents required to discriminate among a set of detailed images is theoretically predicted for the case in which the number of detected photoevents N is fixed. The theoretical predictions are verified experimentally, using the laboratory system described in Section 3.2; excellent agreement is found between theory and experiment. In Section 3.4, the recognition performance of the correlation signal realized using a Poisson-distributed number of detected photoevents is predicted theoretically, and compared to the performance of the fixed N case. (The theoretical behavior of the correlation signal realized using a Poisson-distributed number of photoevents has been previously verified via a Monte Carlo computer simulation¹). Finally, the effects of additive noise on the performance of the recognition system are investigated in Section 3.5. A method for reducing the effects of additive noise is suggested in Section 3.6.

3.2 Quantum-Limited Imaging Systems

When a quantum-limited image is acquired using a position-sensitive, photon-counting detection system, the system's output is a list of the spatial coordinates of the individual detected photoevents. As indicated in Chapter 2, Eq. (2.7), the probability of detecting a photoevent at a given location is directly proportional to the classical intensity of the corresponding location in the input image²². This statistical sampling of the input image provides a natural means of data compression, and provides the basis upon which this technique for image recognition is founded.

Examples of photon-limited images are shown in Fig. 3.1. The photographs in Fig. 3.1 were obtained using the position-sensitive, photon-counting detection system shown schematically in Fig. 3.2. The images in the top row were obtained by detecting the locations of 20 million photoevents, and histogramming the number of detected photoevents at each location into 256 bins. The resulting 256 gray level images were then displayed on a video monitor and photographed. The images realized by detecting $N=4000$ and 1000 are binary.

Most photon-counting imaging systems are based on detectors that utilize a photocathode, a stack of microchannel image intensifiers², and an anode assembly that is used to determine the position of detected photoevents. The main difference between various systems that have been reported is the type of anode assembly that is used. Anode structures that have been reported include self-scanned CCD arrays^{3,4}, grey-coded masks used with a bank of photomultipliers^{5,6}, multi-anode arrays^{7,8}, silicon-intensified-target television cameras^{9,10}, wedge-and-strip anodes¹¹, crossed-wire-grid anodes¹², and resistive anodes¹³⁻¹⁶. More recently, the use of delay-line readout methods^{17,18} have been used in place of an anode assembly, and may prove to be a

superior method for determining the photoevent locations. Using delay-line readout methods, spatial resolutions up to 1000×1000 have been reported, at detection rates up to one million per second¹⁷.

The choice for the anode assembly that is employed in a given detection system depends on the requirements of the particular application. For example, detection systems that detect one photoevent at a time^{3-6, 11-16} can be used to detect a fixed number of detected photoevents, and can provide temporal, as well as positional information about each detected photoevent. The detection systems that utilize a detector array or a vidicon tube to determine the position coordinates are preferable in high speed applications that involve moving objects, or pulsed illumination sources. Detection systems employing these types of anode assemblies can detect multiple photoevents in a short time period, which are then read out in a raster format. Temporal (time of arrival) information about the detected photoevents is lost, however, and the number of detected photoevents in any given photon-limited image is a random variable [see Eqs. (2.5) and (2.6)]

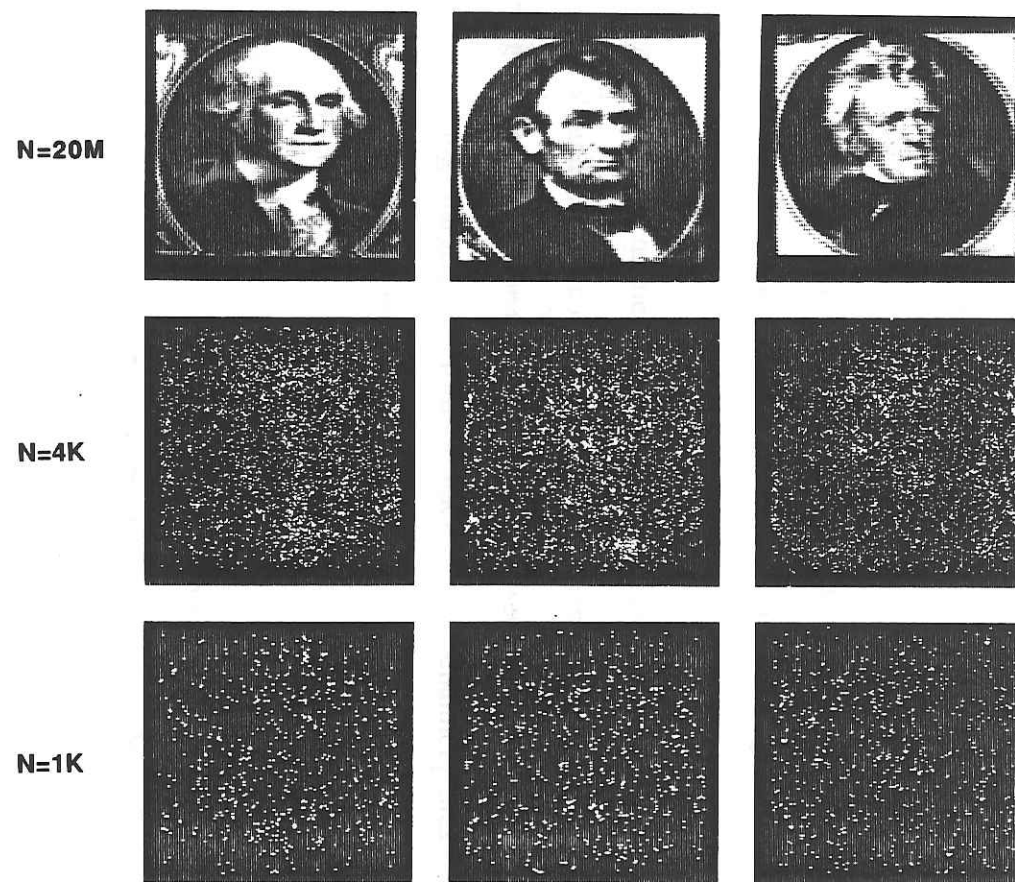


Figure 3.1 Images of engraved portraits obtained using two-dimensional photon-counting detection system: first column, portrait of George Washington; second column, Abraham Lincoln; third column, Andrew Jackson. N is the number of detected photoevents over the entire image. The spatial coordinates of each photoevent are digitized to 8-bit accuracy.

For the experimental work in this thesis, it was desired to have a detection system that allowed one to realize a quantum-limited image with either a fixed or a random number of detected photoevents. This requirement, along with the established commercial availability of the various components of the system, led to the choice of the detection system shown schematically in Fig. 3.2. The operation of this system is described as follows; an incident photon ejects an electron from the photocathode. A potential difference is used to direct the photoelectron onto the stack of microchannel plates. Each microchannel plate consists of a wafer of glass that is permeated by small parallel channels; each channel is lined with a semiconducting material. As a result, each channel acts as a photomultiplier, with the net result being a cascade of electrons from each plate. The channels are set at an angle to enhance gain, and the plates are arranged in V- and Z-stack configurations of the channels to reduce ion feedback. The electron gain provided by the microchannel plate assembly is typically between 10^6 and 10^8 . The resulting charge pulse from the stack of microchannel plates is directed onto the resistive anode, using a potential difference.

The resistive layer covering the anode is terminated by electrodes at four locations around the anode's perimeter. A potential is applied to each electrode, which allows the charge-sensitive preamplifier and position-computing electronics to determine the centroid of the charge distribution, and thereby determine the photoevent location. Using resistive anodes, spatial resolutions as high as 500×500 have been reported¹⁴.

The detection rates at which the resistive-anode based detectors can operate are limited by the speed of the summing amplifiers in the position-computing electronics, and the resolution required for a given application. Detection rates up to 200,000 per

second have been reported¹⁴, with a 128x128 spatial resolution. For eight-bit resolution, detection rates of 100,000 per second are possible^{14,19}.

The dark noise characteristics of the various detection systems depends on the choice of the photocathode. For example, with a bi-alkali (Cs_3Sb) photocathode, which is primarily blue sensitive, the dark count contribution due to thermionic emission of electrons from the cathode is typically less than 100 counts per second at room temperature. It is possible to reduce the dark count rate further by cooling the photocathode. In contrast, a multi-alkali ($\text{Na}_2\text{KSb:Cs}$) photocathode (which is red sensitive) typically produces on the order of 3000 dark counts per second at room temperature. Depending on the particular application, these dark count rates may or may not be important. This issue is addressed in Section 3.5.

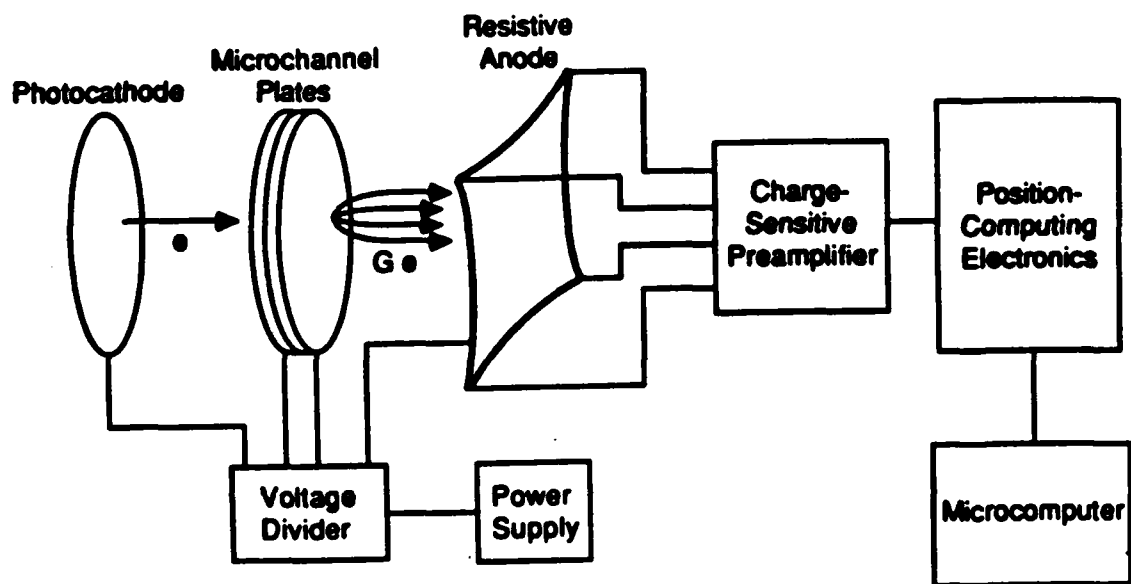


Figure 3.2 Schematic diagram of a resistive-anode based photon-counting detection system.

3.3 Correlation with a Fixed Number of Detected Photoevents

3.3.1 Theoretical Predictions

As mentioned earlier, when a photon-limited input image is cross-correlated with a classical intensity reference image stored in computer memory, the correlation output corresponds to that of a matched filter. Let us first consider the case where the number of detected photoevents N is fixed. Because the reference function is real, the probability density function for the correlation signal is Gaussian, as shown in Eq. (2.56). For fixed N , the mean value and variance are given by Eqs. (2.54) and (2.55) respectively.

The recognition performance was tested on the engraved portraits of George Washington, Abraham Lincoln, and Andrew Jackson shown in Fig. 3.1. The photographs in Fig. (3.1) were obtained using the position-sensitive, photon-counting detection system shown schematically in Fig. 3.2. Portraits from U.S. currency were imaged onto a two-dimensional, photon-counting detector¹⁹ [Electro-Optical Products Div., ITT corporation, Model F4146m]. The detector was connected to position-computing electronics²⁰ to determine the spatial coordinates of the detected photoevents. The spatial coordinates of the detected photoevents were digitized to a spatial resolution of eight bits in each dimension, and then sent to a microcomputer for processing. Illumination was provided by fluorescent room lights, and a neutral density of 4.0 was inserted between the imaging lens and detector to reduce the observed count rate to 50,000 Hz. The detection system has a maximum detection rate of approximately 200 KHz, depending on the spatial resolution desired. The number of detected photoevents for the images in each row is shown at the left of each row.

One calculates the probability density functions for the photon-limited correlation signal given in Eq. (2.9) using the classical intensity images in the top row

of Fig. 3.1. One would expect that a fairly large number of detected photoevents is required for accurate discrimination among these detailed objects. Hence, the normal approximation for the PDF of $C(x,y)$ given in Eq. (2.56) should be valid.

Since the reference and input images have the same area A , we will take the reference window offset to be zero (i.e., the images are registered). For this case of a real reference function and N fixed, the correlation signal $C(x,y)$ is given by

$$C = C(0,0) = \sum_{i=1}^N R(x_i, y_i) \quad . \quad (3.1)$$

Using Eqs. (2.54) and (2.55), the mean value and variance of C are found to be

$$\langle C \rangle = \frac{N \iint_A f(x', y') R(x', y') dx' dy'}{\iint_A f(x', y') dx' dy'} \quad , \quad (3.2)$$

and

$$\sigma^2 = \frac{N \iint_A f(x', y') R^2(x', y') dx' dy'}{\iint_A f(x', y') dx' dy'} - \frac{\langle C \rangle^2}{N} \quad , \quad (3.3)$$

respectively, where $f(x', y')$ denotes the classical-intensity input image, and $R(x', y')$ is the real reference function (in this case the classical-intensity reference object) that is stored in computer memory. The high-light-level image of George Washington [$N=20 \times 10^6$] (see Fig. 3.1) is used as the reference function $R(x', y')$. The number of pixels in the reference image is $N_{\text{pix}} = 128 \times 160 = 20,480$. Based on Eqs. (3.2) and (3.3), Gaussian probability density functions for the correlation signal are calculated; these results are plotted in Fig. 3.3 for the case in which the number of detected photoevents in the entire input image is (a) $N=250$, (b) $N=500$, (c) $N=1000$. Curve I

is the PDF $P(C)$ of the correlation signal C , in Eq. (3.1), when the input image is the portrait of George Washington. Curve II is the PDF when the input is the portrait of Abraham Lincoln; and curve III is the PDF when Andrew Jackson's portrait is input.

As one might expect, when the input scene matches the reference scene, the photon-limited correlation signal tends to be higher (see curves I). However, when the number of detected photoevents is too small [3.3(a)], the statistical spread in the correlation values realized will be large enough that both noise objects sometimes yield correlation values greater than some of the correlation values resulting from the reference object. This is indicated by the overlap in the three probability density functions in Fig. 3.3(a). As more detected photoevents are used to realize the correlation signal, one observes a decrease in the variance of the correlation signal, relative to the expected value of the correlation. This is indicated by the increasing separation of the three curves in Figs. 3.3(b) and 3.3(c). This increase in the separation is clearly in agreement with one's intuition, as one would expect the accuracy of the estimate for the correlation signal to increase when a larger number of photoevents is used to estimate the correlation.

As described in Section 2.5, one employs the theory of hypothesis testing when making a recognition decision based on a single realization of the correlation signal. One chooses some correlation threshold C_T ; if the observed correlation value exceeds the threshold, one decides that the reference object is input. If the observed value is less than C_T , the decision is that a noise object was input. Applying the method described in Section 2.5, the correlation thresholds for each number of detected photoevents are chosen using the theoretical predictions for the probability density function of the correlation signals shown in Fig. 3.3. The correlation threshold C_T is chosen such that the probability of detection is maximized and the probability of false

alarm is minimized. When 250 detected photoevents are used to estimate the correlation signal [see Fig. 3.3(a)], the optimum correlation threshold yields a probability of detection of 0.951, with a probability of false alarm of 0.072. When 500 detected photoevents are used to estimate the correlation signal [Fig. 3.3(b)], the probability of detection is 0.995, while the probability of false alarm is 0.0011. Finally, when 1000 detected photoevents are used to realize the correlation signal, the probability of detection is 0.99996, and the probability of false alarm is 0.00003.

This information is illustrated in Fig. 3.4 in the form of receiver-operator-characteristic (ROC) curves. ROC curves provide a means of graphically displaying the overall detection performance of a recognition system. In a ROC curve, one plots the probability of detection versus the probability of false alarm one would achieve with a given system for increasing values of the correlation threshold C_T . In a perfect system, it would be possible to choose C_T such that the probability of detection would be one, and the probability of false alarm would be zero. Hence, the ROC curve for a perfect system would be a right angle curve, with the vertex in the upper left corner. The performance of a given system can be measured by its deviation from that of a "perfect" system. In Fig 3.4, the probability of detection P_d is plotted versus the probability of false alarm P_{fa} when the image of Washington is the reference function, and the images of Washington and Lincoln are input. Note that as the number of photoevents used to realize the correlation signal increases, the system performance moves closer to that of a "perfect" detection system.

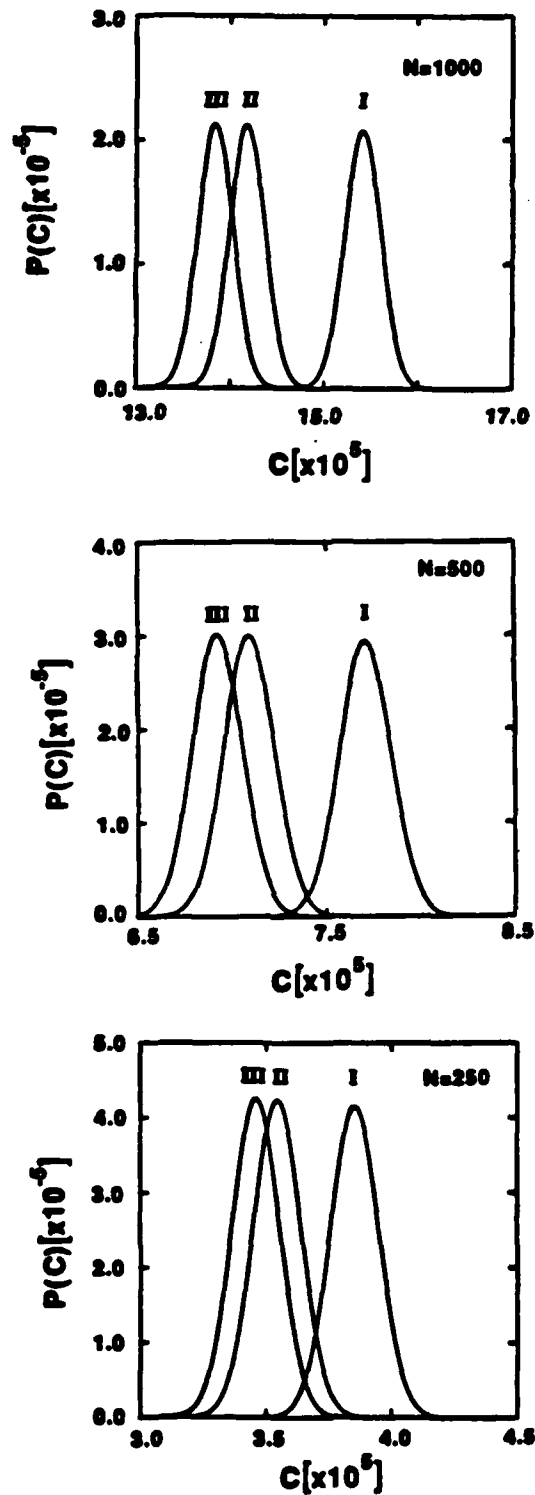


Figure 3.3. Probability density functions of the correlation signal when the input image $f(x',y')$ is the portrait of (I) George Washington, (II) Abraham Lincoln, and (III) Andrew Jackson. The number of detected photoevents is (a) $N=250$, (b) $N=500$, and (c) $N=1000$. The reference function $R(x',y')$ in all cases is the portrait of George Washington with $N=20$ million (see Fig. 3.1).

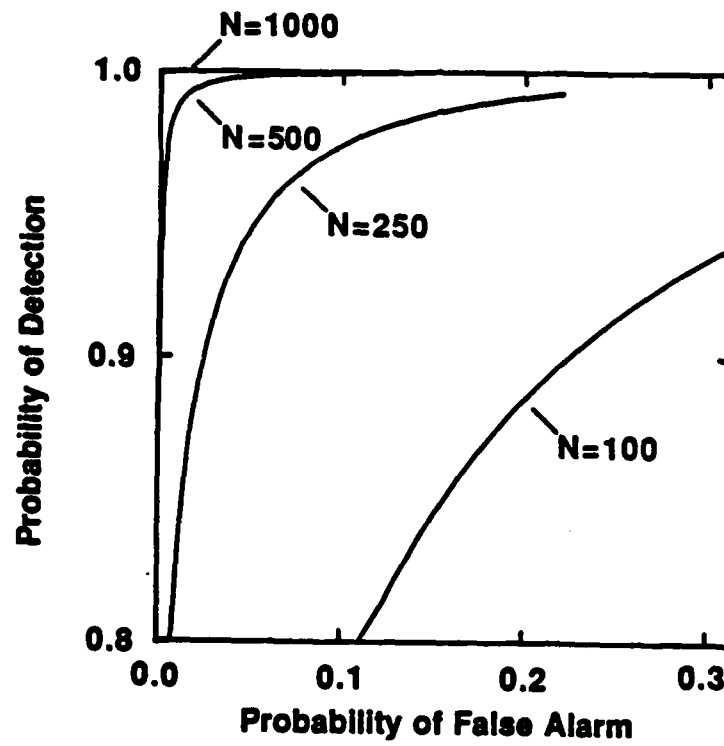


Figure 3.4. ROC curves for the portraits of Washington and Lincoln for different values of N . The reference function $R(x',y')$ in all cases is the portrait of George Washington with $N=20$ million (see Fig. 3.1).

3.3.2 Experimental Results

The theoretical predictions for the behavior of the correlation signal were tested using the experimental configuration shown in Fig. 3.2. One thousand realizations of the correlation signal were made for each input image. This was performed for the number of detected photoevents $N=1000$, 1500, and 2000. (The number of detected photoevents were chosen to obtain a low error rate). A tabular comparison of the agreement between theory and experiment is given in Table 3.1 for each number of detected photoevents. The results for the case in which $N=1000$ is shown in histogram form in Fig. 3.5. One thousand realizations of the correlation signal were performed for each input image; the image of Washington was the reference in all cases. In Fig. 3.5, the solid lines are theoretical predictions for the density function of the correlation signal when the input object is (I) Washington, (II) Lincoln, and (III) Jackson. In Fig. 3.5, one sees excellent agreement between theoretical predictions (solid curves) and the laboratory measurements (histograms). It is important to note that no corrections were made for additive noise effects or dead-time effects. These effects are simply not important when the input count rate is 50,000 Hz., and N is at least several hundred detected photoevents. In the experiments, the dark count rate was observed to be approximately 50 Hz. Hence, at a rate of 50,000 Hz., on average only one detected photoevent out of a thousand is associated with additive noise. As a result, the contribution due to additive noise is clearly negligible. For a discussion of the effects of additive noise on the correlation signal, the reader is directed to Sec. 3.5

Input:		Washington		Lincoln		Jackson	
		Mean	Sigma	Mean	Sigma	Mean	Sigma
N=1000	Theory	1.54E6	1.92E4	1.42E6	1.88E4	1.38E6	1.88E4
	Exp	1.54E6	1.95E4	1.42E6	1.93E4	1.38E6	1.95E4
N=1500	Theory	2.31E6	2.35E4	2.13E6	2.32E4	2.08E6	2.30E4
	Exp	2.31E6	2.48E4	2.13E6	2.29E4	2.07E6	2.24E4
N=2000	Theory	3.08E6	2.72E4	2.84E6	2.66E4	2.77E6	2.65E4
	Exp	3.08E6	2.68E4	2.84E6	2.71E4	2.77E6	2.71E4

Table 3.1 Comparison of theoretical predictions and experimental results for the photon-limited correlation signal realized using various numbers of detected photoevents. In each case, the reference image was the image of Washington (N=20 million) shown in Fig. 3.1.

Image Correlation

Ref: Washington

N=1000

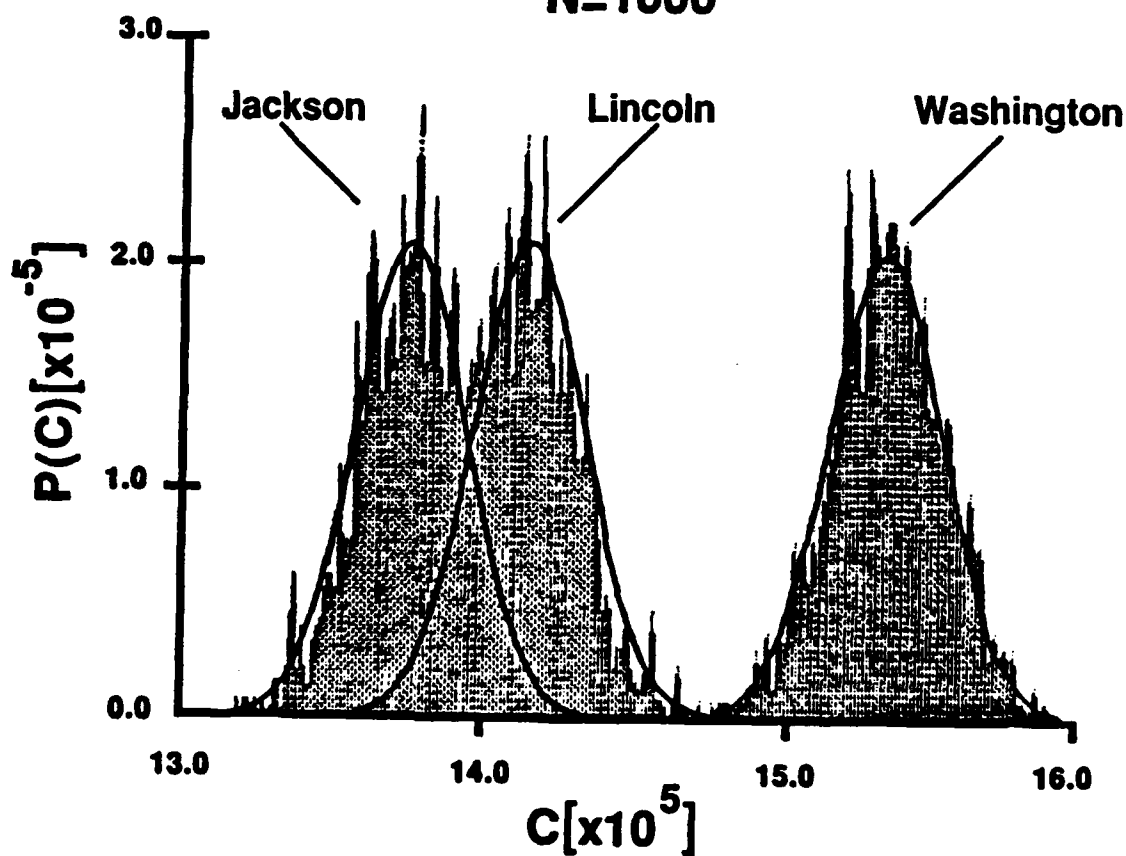


Figure 3.5 Histogram of correlation values obtained from laboratory measurements of the photon-limited correlation signal when the input objects $f(x',y')$ were the portraits of Washington, Lincoln and Jackson. In each realization the total number of detected photoevents was $N=1000$ (see bottom row of Fig. 3.1). The reference function was the portrait of Washington with $N=20$ million.

3.4 N Poisson Distributed

As mentioned earlier, it is possible to perform photon-limited image correlation in two ways. One way is to detect a fixed *number* of photoevents, and use that fixed number of photoevents to compute the quantum-limited correlation signal. This is what is described in the preceding section. The second way of performing the image correlation is to detect photoevents for a fixed *time*; this makes the number of detected photoevents a random variable. In the second method there is one more source of randomness in any realization of the correlation signal, which makes the variance in the estimate of the correlation larger. This was explained theoretically in Section 2.3. In this section, as a means of comparing the effectiveness of the two methods, theoretical predictions for the photon-limited correlation signals are plotted for the same input and reference images that were used in the previous section.

For convenience, we repeat the equations for the mean value and variance of the photon-limited correlation signal, when N is Poisson-distributed. The probability density function is again normal, with mean value and variance given by (see Eqs. (2.54), (2.55) and (2.7))

$$\langle C \rangle = \frac{\bar{N} \iint_A f(x', y') R(x', y') dx' dy'}{\iint_A f(x', y') dx' dy'} \quad , \quad (3.4)$$

and

$$\sigma^2 = \frac{\bar{N} \iint_A f(x', y') R^2(x', y') dx' dy'}{\iint_A f(x', y') dx' dy'} \quad , \quad (3.5)$$

respectively, where once again $f(x', y')$ denotes the classical-intensity input image, and $R(x', y')$ is the real reference function. In Fig. 3.6, theoretical predictions for the

probability density functions of the correlation signal are plotted when the reference function is the image of Washington, and the input images are again (I) Washington, (II) Lincoln, and (III) Jackson.

The mean number of detected photoevents \bar{N} was chosen to roughly achieve the same relative separations that occurred between the density functions in Fig. 3.3, for the fixed N case. In Fig. 3.7, ROC curves are plotted for the case when the number of photoevents is Poisson distributed, and the input images are Washington and Lincoln. Note that almost three times as many detected photoevents are required to achieve the same relative probabilities of detection and false alarm that were achieved in the fixed N case (see Fig. 3.4). This is due to the increased variance observed in the correlation signal when the number of detected photoevents is random [see Eqs. (2.52) and (2.55)].

At this point, it is instructive to address the issue of why, given the choice, would one perform the photon-limited image correlation experiments using a Poisson-distributed number of detected photoevents, when detecting a fixed number provides the same results using fewer photoevents. Clearly, given the choice, detecting a fixed number of photoevents is the superior method. However, there may be cases in which it is not possible to detect a fixed number of photoevents. One important example of this is the case in which the input to the photon-counting recognition system is a video monitor, or some other input device that operates via a raster effect. In this case, one must detect photons for an *integral number of frame times*, rather than detect a fixed number of photoevents. The justification for this is the following. Consider the case in which one wishes to detect, say, 5000 photoevents, and the detection rate is 50,000 Hz. It will take on average 10 milliseconds to perform one realization of the correlation signal. Note, however, that this is only one third of the frame time of a standard video

monitor. As a result, the raster only scanned one third of its entire pattern, so the entire image on the video monitor was not presented to the photon-counting detection system. Now, if one detects photoevents for an integral number of frame times, then the raster pattern will present all areas of the image from the video monitor equally, and no further synchronization is necessary. Additional applications in which it is necessary to detect photoevents for a fixed time include any experiments involving the detection of pulsed light.

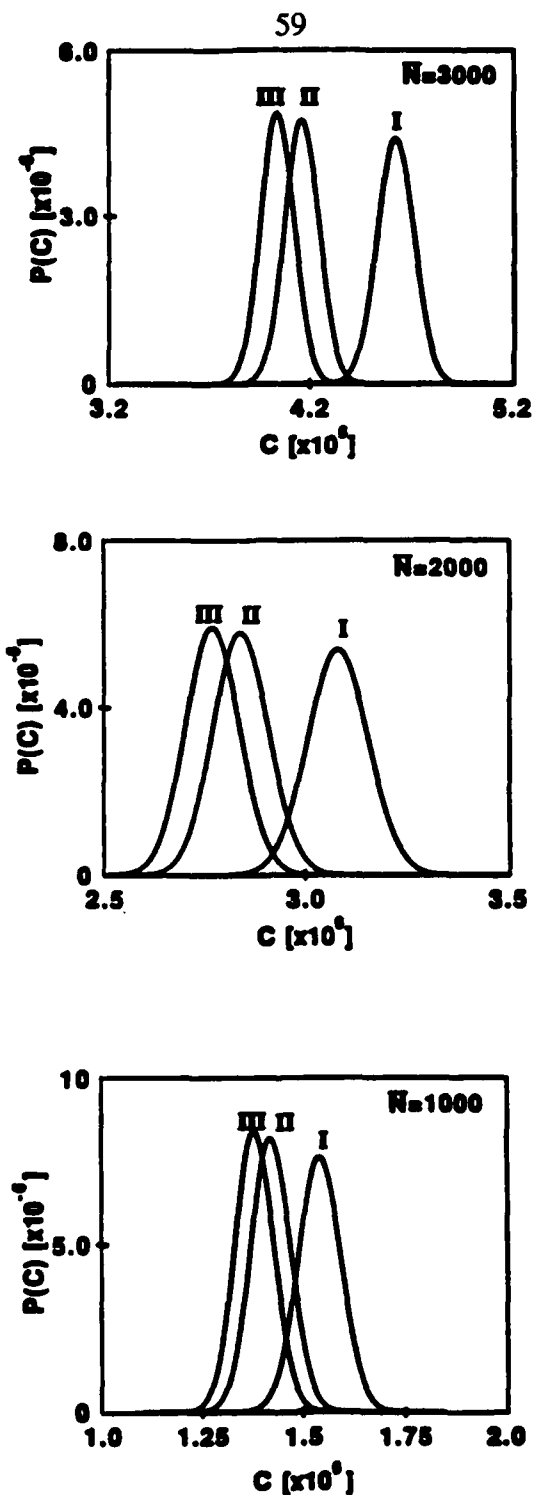


Figure 3.6 Probability density functions of the correlation signal when the input image $f(x',y')$ is the portrait of (I) George Washington, (II) Abraham Lincoln, (III) Andrew Jackson. The average number of detected photoevents is (a) $\bar{N}=1000$, (b) $\bar{N}=2000$, and (c) $\bar{N}=3000$. The reference function $R(x',y')$ is the portrait of Washington.

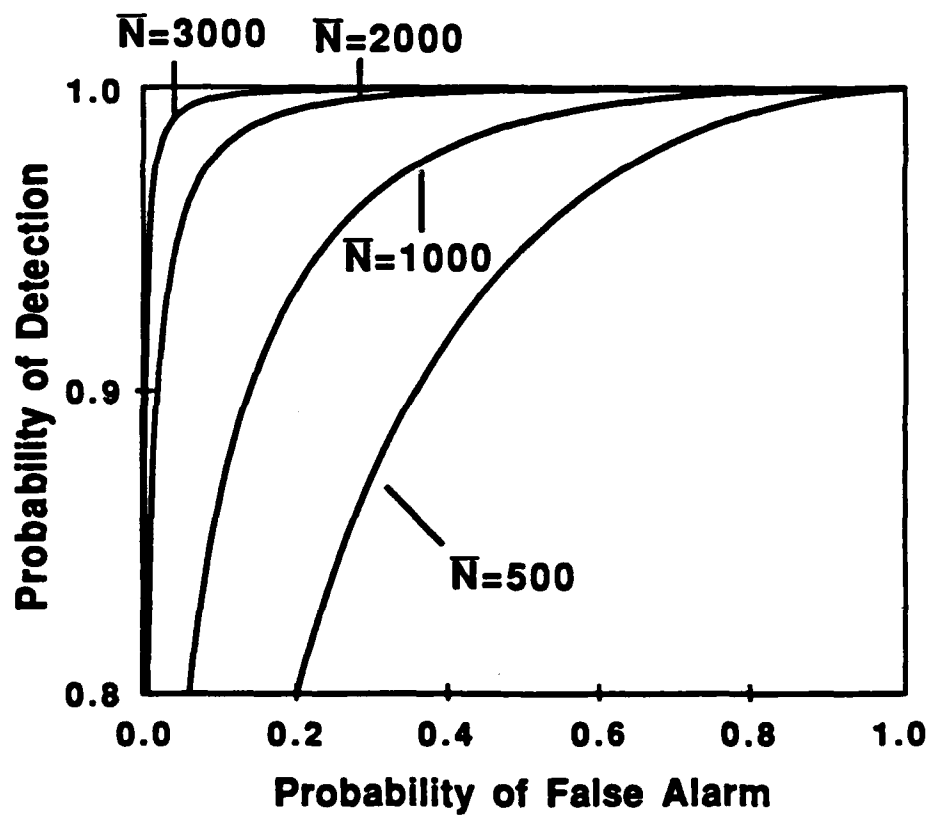


Figure 3.7 ROC curves for Washington and Lincoln for various values of \bar{N} .

3.5 Effects of Additive Noise

3.5.1 Introduction

In two-dimensional, photon-counting detectors, the dominant source of additive noise is associated with dark counts, due primarily to thermal emission of electrons from the photocathode. Here, we consider the effects of additive noise on the correlation signal for the case in which the number of detected photoevents is Poisson-distributed, and the reference function is real.

The number of dark counts actually observed in an experimental realization of a photon-limited correlation signal depends on the dark count detection rate, R_D , the total photoevent detection rate, R_T , and the number of detected photoevents, $N=N_f+N_D$, used to realize the correlation signal. Here, N_f denotes the number of detected photoevents resulting from the input image $f(x,y)$, and N_D denotes the number of dark counts in the particular realization of the correlation signal. For example, if $R_T \gg R_D$, and N is large (say a few hundred), then the number of dark counts observed in any given realization of the correlation signal will be small. This is the case for all of the image recognition experiments described in the thesis. Typical experimental values are $R_T = 50$ KHz, $R_D=50$ Hz., and $N=1000$. Given these values, the number of dark counts actually observed in any realization of the correlation signal is approximately one, on average. Obviously, in this case the effect of the dark noise is not significant.

The effect of the dark noise becomes significant when the input illumination is insufficient to achieve total detection rates such that $R_T \gg R_D$. In this section, the effect of additive noise on the recognition performance of the photon-limited correlation signal

is considered for various amounts of dark noise. To facilitate comparison, theoretical predictions are made using the engraved portrait images shown in Fig. 3.1.

3.5.2 Effect on Recognition Performance

From Eq. (2.9), the total photon-limited correlation signal can be written as²¹

$$C(x, y) = \sum_{i=1}^{N_f} R(x + x_i, y + y_i) + \sum_{j=1}^{N_D} R(x + x_j, y + y_j) \quad , \quad (3.6)$$

where N_f is the number of detected photoevents due to the input scene $f(x', y')$, and N_D is the number of dark counts observed. If photoevents are detected for a fixed time, then both N_f and N_D are Poisson-distributed, and statistically independent²². By applying the same analysis that was performed in Chapter 2, it readily follows that the probability density function for the photon-limited correlation signal will again be normal,

$$P(C(x, y)) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \left\{ \frac{(C(x, y) - \langle C(x, y) \rangle)^2}{2\sigma^2} \right\} \quad . \quad (3.7)$$

The mean value of the correlation signal is

$$\begin{aligned} \langle C(x, y) \rangle = & \bar{N}_f \iint_A p[(x', y') | f(x', y')] R(x + x', y + y') dx' dy' \\ & + \bar{N}_D \iint_A p[(x', y') | D(x', y')] R(x + x', y + y') dx' dy' \quad , \end{aligned} \quad (3.8)$$

and the variance is

$$\begin{aligned} \sigma^2 = & \bar{N}_f \iint_A p[(x',y')|f(x',y')] R^2(x+x',y+y') dx' dy' \\ & + \bar{N}_D \iint_A p[(x',y')|D(x',y')] R^2(x+x',y+y') dx' dy' . \end{aligned} \quad (3.9)$$

In Eqs. (3.8) and (3.9), $p[(x',y')|f(x',y')]$ is given by Eq. (2.7). The spatial distribution of the dark counts will depend on the characteristics of the individual detection system, and is denoted by $D(x',y')$. The probability of detecting a dark count, given the spatial distribution of the dark counts, is given by

$$p[(x',y')|D(x',y')] = \frac{D(x',y')}{\iint_A D(x',y') dx' dy'} \quad (3.10)$$

Note that the presence of additive noise shifts the mean of the observed correlation signal, and increases the variance.

In Fig. 3.8, the probability density functions for the case in which Washington is the reference, and (I) Washington, (II) Lincoln and (III) Jackson are again input, using $\bar{N}=3000$ detected photoevents. Note that this is the number of detected photoevents that provided a probability of detection of 0.982, with a probability of false alarm equal to 0.031 (see Fig. 3.7). In this case, various amounts of additive noise were taken to be present in the correlation signal. In Fig. 3.8 , top, 10% noise was present (i.e., 2700 signal photons, and 300 noise photons), in Fig. 3.8, middle, 20%, and in Fig. 3.8, bottom, 30%. In Fig. 3.9, the ROC curves are plotted for various amounts of additive noise. Once again, the reference is the image of Washington, and the inputs are Washington and Lincoln. The probability density of the dark counts $P[(x',y')|D(x',y')]$ was determined by operating the photon-counting

detection system in total dark darkness for 12 hours, during which time 2.1 million dark counts were detected. It is important to note that the recognition performance degrades relatively rapidly with increasing amounts of noise. A method for reducing the effects of the additive noise is given in the next section.

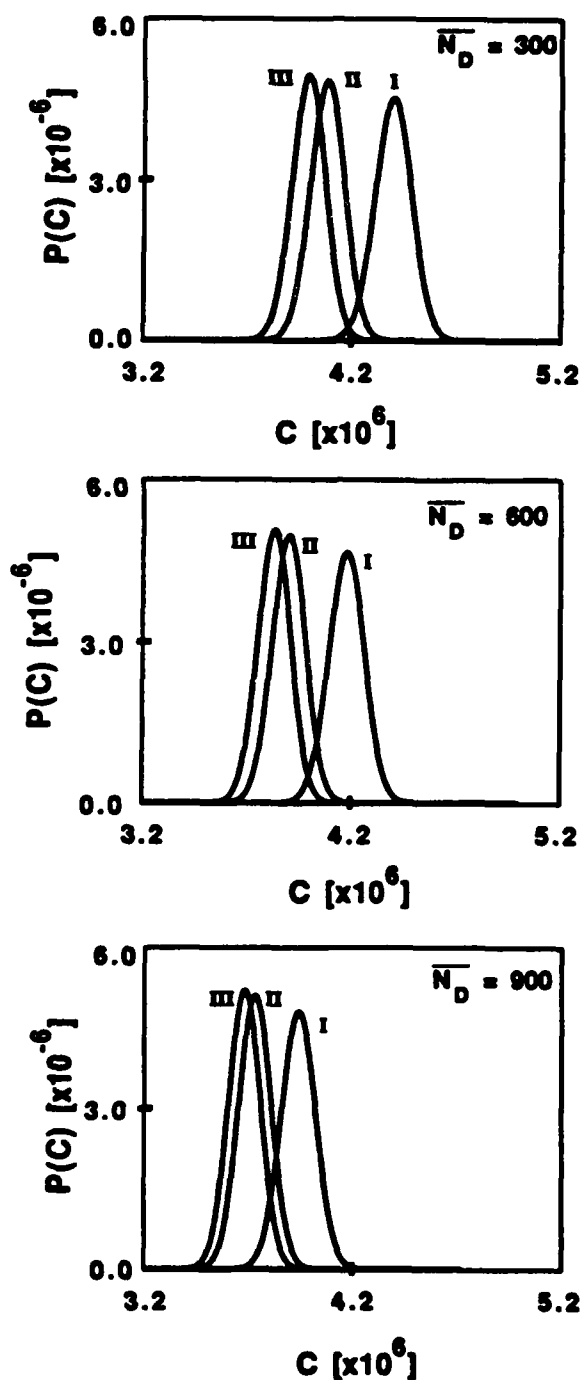


Figure 3.8 Effect of dark noise on the probability density functions of the correlation signal. The input image $f(x',y')$ is the portrait of (I) George Washington, (II) Abraham Lincoln, (III) Andrew Jackson. In all cases, the total average number of detected photoevents is $\bar{N}=3000$. The reference function $R(x',y')$ in all cases is the portrait of George Washington.

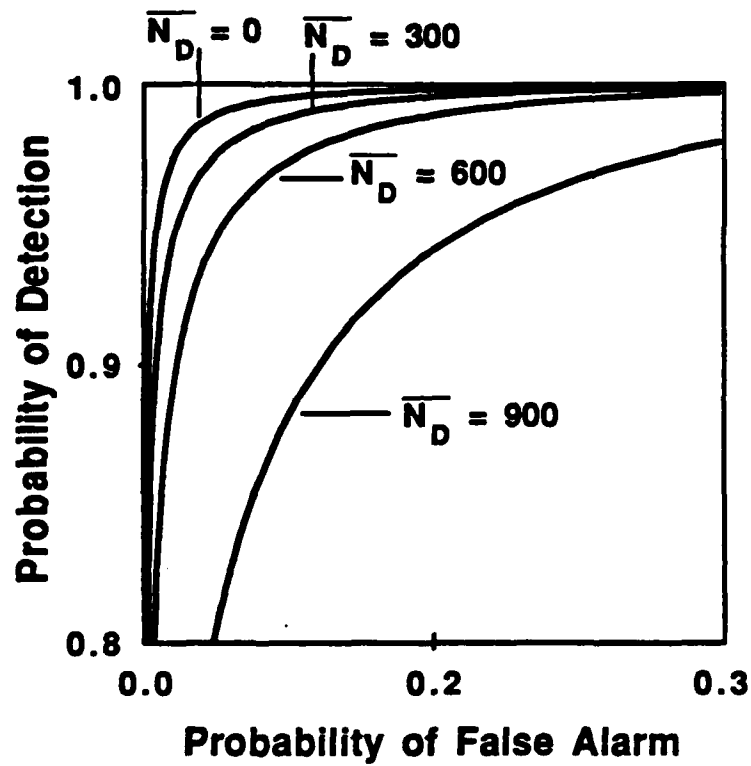


Figure 3.9 ROC curves for Washington and Lincoln in the presence of various amounts of additive noise \bar{N}_D . In all cases, the total average number of detected photoevents is $\bar{N}=3000$.

3.5.3 Minimizing the Effects of Additive Noise

It is important to note that the effect of the dark noise is independent of the input image; it depends only upon the individual characteristics of the photon-counting detector, and the reference function that is chosen for a particular realization of the correlation signal. Hence, the effects of the dark counts can be reduced. For example, consider the case in which the additive noise accounts for 50% (on average) of the total detected photoevents, and it is determined that an average number of 3000 photoevents is required to accurately recognize the input image in the absence of noise (as is the case for the images of the presidents used in this chapter, when N is Poisson-distributed). To reduce the effects of the dark noise, one would do the following. First, one would recognize that the correlation signal must be realized using $\bar{N}=6000$ detected photoevents, as only 3000 (on average) of those result from the input signal. Next, one computes the bias that must be subtracted from the experimental measurement of the photon-limited correlation signal. This bias is given by the second term in Eq. (3.6). Note that this bias is different for each reference function. Subtraction of the bias will reduce the effects of the noise on the mean value of the correlation signal, but it does not reduce the variance. Hence, it is not possible to eliminate the effects of the dark noise entirely, but the effects can be greatly reduced.

If one is to use this bias subtraction method in recognition experiments, one must obtain theoretical expressions for the probability density function, mean value, and variance of the "corrected" correlation signal. In this case, the corrected correlation signal $C'(x,y)$ can be written as

$$C'(x,y) = \sum_{i=1}^{N_I} R(x+x_i, y+y_i) + \sum_{j=1}^{N_D} R(x+x_j, y+y_j) - B_D(x,y) \quad , \quad (3.11)$$

where the bias that is subtracted, $B_D(x,y)$, is given by

$$B_D(x, y) = \bar{N}_D \iint_A p[(x', y') | D(x', y')] R(x + x', y + y') dx' dy' \quad (3.12)$$

The probability density function is obtained from the "uncorrected" PDF by making the following change of variables. Let $C'(x, y) = C(x, y) - B_D$. It follows directly that the PDF for $C'(x, y)$ is given by

$$P(C'(x, y)) = \frac{1}{\sqrt{2\pi}\sigma'} \exp - \left\{ \frac{(C'(x, y) - \langle C'(x, y) \rangle)^2}{2\sigma'^2} \right\} \quad (3.13)$$

where the mean value $\langle C'(x, y) \rangle$ is given by

$$\langle C' \rangle = \frac{\bar{N}_f \iint_A f(x', y') R(x', y') dx' dy'}{\iint_A f(x', y') dx' dy'} \quad (3.14)$$

and the variance σ'^2 is given by Eq. (3.9). Note that the bias subtraction has eliminated the shift in the mean value, but did not correct the increase in the variance.

In Fig. 3.10, ROC curves are plotted for various amounts of additive noise. Once again, the reference is the image of Washington, and the inputs are Washington and Lincoln. Note that the performance does not degrade as rapidly with increasing amounts of noise, when the bias term $B_D(x, y)$ [see Eq. (3.12)] is subtracted. Comparing Fig. 3.10 with Fig. 3.9, it is important to note that the effect of the bias subtraction is greatest when the amount of noise is large (e.g. $\bar{N}_D = 900$.) This is indicated by the fact that the ROC curve in Fig. 3.9 for the case when $\bar{N}_D = 900$ comes closer to the upper left hand corner of the graph than the corresponding curve in Fig. 3.8. In other words, the probability of detection increases at a faster rate with the bias

subtracted. Comparing the curves for $\bar{N}_D=300$ and $\bar{N}_D=600$ in Figs. 3.8 and 3.9, the improvement due to the bias subtraction is less pronounced, as one might expect.

In practice, to obtain the best recognition performance in the presence of a known amount of additive noise, one would subtract the appropriate bias, given in Eq. (3.12) from experimental realizations of the correlation signal. Theoretical predictions for the recognition performance can be made using Eqs (3.9), (3.13) and (3.14).

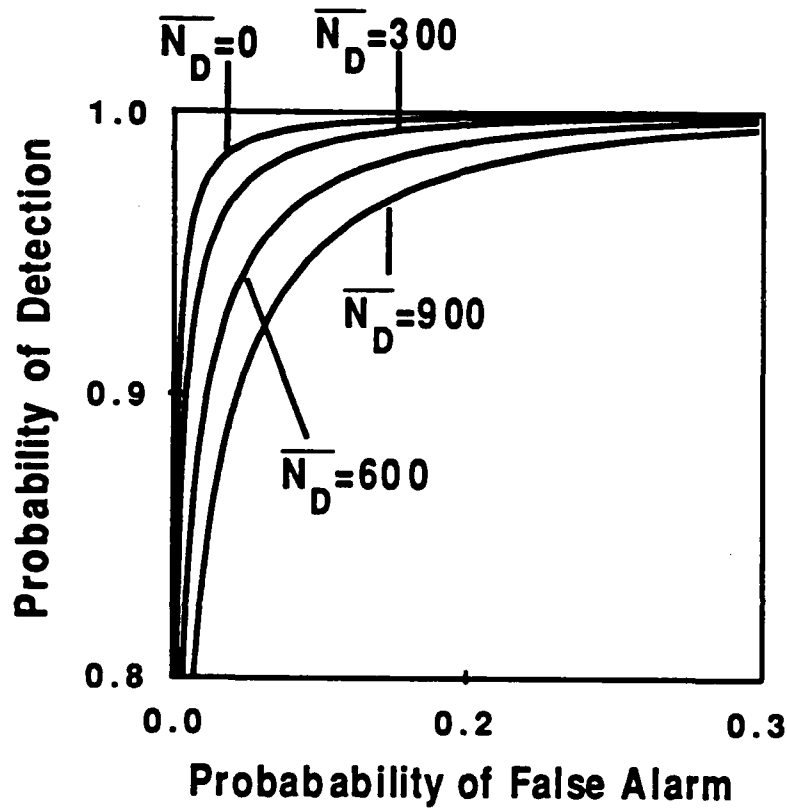


Figure 3.10 ROC curves for Washington and Lincoln in the presence of various amounts of additive noise, with the bias subtracted. In all cases, the total average number of detected photoevents is $\overline{N}=3000$.

3.6 Summary

In this chapter, the behavior of the correlation signal, realized by cross-correlating a photon-limited input image with a classical-intensity reference image, is analyzed for various cases. In Section 3.3.1, theoretical predictions are made for the behavior and recognition performance of the photon-limited correlation signal realized using a fixed number of detected photoevents. The images used in the recognition experiments are shown in Fig. 3.1. The theoretical predictions are verified experimentally in Section 3.3.2, where excellent agreement is obtained between theory and experiment (see Fig. 3.5). In Section 3.4 the behavior of the correlation signal realized using a Poisson-distributed number of detected photoevents is analyzed using the same input images. The recognition performance is compared to the case in which the number of detected photoevents is fixed; the fixed-N method is found to yield superior results. Finally, in Section 3.5.1, the effect of additive noise on the recognition performance of the photon-limited correlation signal is analyzed. Theoretical predictions are plotted in Figs. 3.8 and 3.9. In Section 3.5.2, a method for reducing the effects of additive noise is described, with the results plotted in Fig. 3.10.

References for Chapter 3

1. G. M. Morris, 'Scene matching at low light levels: a computer simulation,' *Appl. Opt.* **23**, 3152-3159 (1984).
2. M. Lampton, 'The microchannel image intensifier,' *Sci. Am.* **245**, 62-71 (1981).
3. C. B. Johnson and R. E. Blank, 'Image tube intensified linear and area self-scanned array detectors for astronomy,' *Proc SPIE* **290**, 102-108 (1981).
4. E. Roberts, T. Stapinski and A. Rodgers, 'Photon-counting array detectors for the Fuse/Lyman satellite telescope,' *J. Opt. Soc. Am. A* **3**, 2146-2150 (1986).
5. C. Papaliolios and L. Mertz, 'New two-dimensional photon camera,' *Proc SPIE* **331**, 360-364 (1982).
6. C. Papaliolios, P. Nelson and S. Ebstein, 'Speckle imaging with the PAPA Detector,' *Appl. Opt.* **24**, 287-292 (1985).
7. J. G. Timothy, G. H. Mount and R. L. Bybee, 'Detector arrays for photometric measurements at soft x-ray, ultraviolet and visible wavelengths,' *Proc SPIE* **183**, 169-181 (1979).
8. J. G. Timothy, 'Multi-anode microchannel array detector systems: performance characteristics,' *Opt. Eng.* **24**, 1066-1071 (1985).
9. A. Blazit, L. Koenchlin and J. L. Oneto, 'On line digital correlation of photon-counting TV images for stellar interferometry,' in *Image Processing Techniques in Astronomy*, C. de Jager and H. Nieuwenhuijzen, eds. (D. Reidel, Dordrecht-Holland, 1975) p. 79-84.
10. P. B. Boyce, 'Low light detectors for astronomy,' *Science* **198**, 145-148 (1977).
11. H. S. Schwartz and J. S. Lapington, 'Optimization of wedge and strip anodes,' *IEEE Trans. Nucl. Sci.* **NS-32**, 433-437 (1979).
12. E. M. Kellogg, S. S. Murray, and D. Bardas, *IEEE Trans. Nucl. Sci.* **NS-26**, 403-410 (1979).
13. M. Lampton and C. W. Carlson, 'Low-distortion resistive anodes for two-dimensional position-sensitive MCP systems,' *Rev. Sci. Instrum.* **50**, 1093-1097 (1979).

14. M. Lampton and C. W. Carlson, 'Low-distortion resistive anodes for two-dimensional position-sensitive MCP systems,' *Rev. Sci. Instrum. Eng.* **331**, 365-371 (1982).
15. A. H. Greenaway, A. Lyons, I. Mcwerter, D. Rees and A. Cochrane, 'Miniature imaging photon detectors III. An assessment of the performance of the resistive anode IPD,' *J. Phys. E* **15**, 145-150 (1982).
16. T. Gonsiorowski, 'Variable threshold discrimination in a photon-imaging detector,' *Appl. Opt.* **23**, 1060-1065 (1984).
17. S. E. Sobattka and M. B. Williams, 'Delay line readout of microchannel plates,' *IEEE Trns. Nucl. Sci.* **35**, 348-351 (1988).
18. J. A. Harder, 'A fast time-to-digital converter for position-sensitive radiation detectors with delay line readouts,' *Nucl Instrum. Meth. Phys. Res.* **a265**, 500-510 (1988).
19. ITT Electro-Optical Products Division, Tube and Sensors Laboratories, Fort Wayne, IN. Detector model F4146M.
20. Surface Science Laboratories Inc., Mountain View, CA. Model 2401 position computer.
21. G. M. Morris, M. N. Wernick and T. A. Isberg, 'Image correlation at low light levels,' *Opt. Lett.* **10**, 315-317 (1985).
22. L. Mandel, E. C. G. Sudarshan and E. Wolf, 'Theory of photoelectric detection of light fluctuations,' *Proc. Phys. Soc.* **84**, 435-444 (1964).

Chapter 4

Rotation-Invariant Image Recognition at Low Light Levels

4.1 Introduction

The earliest attempts at performing rotation-invariant image recognition involved the use of a multiplexed frequency plane filter, which was initially suggested by Vander Lugt¹. The multiplexed filter is produced by recording several matched filters at the same spatial location in the Fourier transform plane, where each filter is matched to a different in-plane rotational orientation of the reference object. Clearly, the holographic implementation of this method becomes cumbersome for complex objects, since many orientations must be recorded to achieve satisfactory recognition performance. For example, Casasent and Psaltis² reported that when a reference object was rotated by 3.5 degrees, or changed scale by 2% with respect to the input, the output signal-to-noise ratio (SNR) of the correlation peak dropped from 30 dB to 3 dB. (The rate of decrease in SNR increases with the space-bandwidth product of the image). These results imply that as many as 360 orientations must be recorded, as well as a large number of different sizes. While this technique is difficult to perform holographically, a digital implementation may be practical at some point in the future when many megabytes of memory are cheaply available for standard digital processors.

An elegant solution to the problem of rotation-invariant image recognition is the method of rotation-invariant filtering. A rotation-invariant filter is a correlation filter whose correlation output remains unchanged when the input function undergoes an in-plane rotation. In addition, the correlation output should be a maximum when a

reference function that is to be identified is input. Various types of rotation-invariant filters have been suggested. These include the complex circular-harmonic filter³⁻⁸, the "lock and tumbler" filter^{9,10}, synthetic discriminant functions¹¹⁻²¹ and phase-only filters²²⁻²⁴. More recently, neural networks²⁵ have been employed to perform rotation-invariant pattern recognition and pattern classification²⁶.

The theoretical formalism presented in Chap. 2 can describe the behavior of all of the above filters at low light levels. For this work in rotation-invariant image recognition at low light levels, a correlation filter was selected that required no pre-processing of the input scene. The filter that was most suited to this work was the rotation-invariant circular-harmonic filter. The circular-harmonic filter has been shown to be effective in gray-level images, and can provide information about the orientation of the reference object³⁻⁸. The lock and tumbler filter^{9,10} has primarily been used with binarized images, and synthetic discriminant functions¹¹⁻²⁰ are used primarily for image classification, as opposed to image recognition. Because this is a digital implementation of the correlation filtering, there may be no advantage to the use of a phase-only correlation filter. The phase-only filter's primary advantage is realized in optical implementations, where the phase-only filter has a much higher diffraction efficiency than a phase and amplitude filter²². The theoretical formalism presented here, and in Chapter 2 can, however, predict the behavior of all of the above filters when implemented at low light levels.

The rotation-invariant circular-harmonic filter is briefly reviewed in Section 4.2. In Section 4.3, the theory required to analyze the performance of the circular-harmonic filter at low light levels is presented. Experimental confirmation of the theory detailed in Section 4.3 is presented in Section 4.4. Rotation-invariant recognition from within a

cluttered environment is addressed in Section 4.5, and some additional issues are discussed in Section 4.6

4.2 Rotation-Invariant Filtering using Circular-Harmonic Functions

Any two-dimensional function $f(r, \theta)$ can be represented in terms of its circular-harmonic components as follows:

$$f(r, \theta) = \sum_{m=-\infty}^{\infty} F_m(r, \theta) \exp(im\theta) \quad , \quad (4.1)$$

where

$$F_m(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) \exp(-im\theta) d\theta \quad . \quad (4.2)$$

In Eq. (4.1), $F_m(r, \theta)$ is known as the m^{th} circular-harmonic component of the function $f(r, \theta)$. Other investigations have shown³⁻⁸ that rotation-invariant filtering can be achieved by taking the reference function $R_m(r, \theta)$ to be the complex conjugate of a single (or multiple) circular-harmonic component(s) of the reference object $R(r, \theta)$, e.g.,

$$R(r, \theta) = R_m^*(r) \exp(-im\theta) \quad , \quad (4.3)$$

where $R_m(r)$ is given by Eq. (4.2). The cross-correlation of an input function $f(r, \theta)$ with $R_m(r, \theta)$ is given by

$$C(r, 0) = \int_0^{2\pi} \int_0^{\infty} R_m^*(r + r', \theta') f(r', \theta') r' dr' d\theta' \quad . \quad (4.4)$$

Using the expansion in Eq. (4.1) for $f(r,\theta)$ and substituting for $f(r,\theta)$ in Eq. (4.4) yields

$$C(r,0) = 2\pi \int_0^{\infty} R_m^*(r+r') F_m(r') r' dr' \quad , \quad (4.5)$$

which, when properly normalized attains its maximum when $R_m(r)=F_m(r)$. (The normalization issue is addressed in Sec. 4.5). Note that only the m^{th} circular-harmonic component of the input contributes to the correlation output; this is the result of the orthogonality of the theta integration. In addition, the limits of the r integration are limited to the area of the reference object. If the input function is rotated by an angle α with respect to the reference, the correlation signal becomes

$$C(r,\alpha) = C(r,0)\exp(im\alpha) \quad . \quad (4.6)$$

The effect of rotating the input by an angle α with respect to the reference is contained entirely in the phase term $\exp\{-im\alpha\}$. Therefore, the squared modulus of the correlation signal is independent of the rotation angle α ; i.e.

$$|C(r,\alpha)|^2 = |C(r,0)|^2 \quad . \quad (4.7)$$

independent of the orientation angle alpha. If the reference image is input, then $G_m(r)=F_m(r)$, and the integral in Eq. (4.5) is real-valued. Hence, the orientation of the reference can be obtained from the phase of the correlation signal. The orientation angle α is given by

$$\alpha = \left[\frac{1}{m} \right] \tan^{-1} \left[\frac{\text{Im}\{C(r,\alpha)\}}{\text{Re}\{C(r,\alpha)\}} \right] \quad , \quad (4.8)$$

which can be easily calculated when the correlation is implemented digitally. The rotation angle α is far more difficult to obtain in an optical implementation, because most detection systems are intensity-based.

The magnitude of the correlation signal in Eq. (4.7) will depend on which order harmonic is chosen as the reference function. The optimum circular-harmonic reference function can be determined using the Hotelling trace criterion*. Also, it is evident from Eq. (4.7) that the magnitude of the correlation signal, though independent of the orientation, is dependent on the offset position, r .

Moreover, the magnitude of $C(r, \alpha)$ is dependent on the location of the point in the reference image about which the reference circular harmonic was expanded; this location is referred to as the expansion center⁴. By definition, the optimum (or proper) center is the expansion center that produces the highest autocorrelation peak when the reference image is cross-correlated with the reference circular harmonic. To identify the optimum expansion center, one must perform an exhaustive search of all possible locations in the reference image. However, one does not need to compute an entire circular-harmonic component for each possible expansion center. Using Eq. (4.5), one can see that the output correlation signal is given by the cross-correlation of the radial parts of the circular harmonics in question. Hence, it is sufficient to use Eq. (4.2) to compute $R_m(r)$ by expanding about each point on the reference image. The expansion center that will produce the highest autocorrelation peak is identified by computing the value of

$$\int_0^{\bar{r}} R_m^*(r') R_m(r') dr' \quad , \quad (4.9)$$

for each expansion center location. The point where Eq. (4.9) achieves its maximum value is then the optimum choice for the expansion center of the reference harmonic.

The location of the proper center may be different for each order circular harmonic. In many cases, when the centroid of the input object is used as the expansion center, the resulting autocorrelation peak is sufficient to provide accurate rotation-invariant recognition^{4,6,27}. This is particularly true for applications in which a single object is identified within an uncluttered scene. However, the goal is to locate an object from within a cluttered scene, it is imperative that the proper center be used⁶. In addition, it may be necessary to employ proper normalization [such as the Schwarz inequality, (see Eq. (1.1))] of the correlation signal when operating in a cluttered environment. Inherent limitations often make this difficult in standard optical and digital implementations of the cross-correlation; in the low-light-level implementation it is possible, at least in part, to perform a normalization in real time. This is discussed further in Sec. 4.5.

4.3 Rotation-Invariant Filtering at Low Light Levels

Rotation-invariant filtering at low light levels is achieved by cross correlating a photon-limited input scene $f(r, \theta)$ with the complex conjugate of the m^{th} circular-harmonic component of a reference image $R(r, \theta)$. The mean value of the correlation signal, $\langle C \rangle = \langle C' \rangle + i \langle C'' \rangle$ is given in polar coordinates by (see Eqs. (2.43)-(2.44), with Eqs. (4.3) and (4.6))

$$\langle C(r, \alpha) \rangle = \frac{2\pi \bar{N} \exp\{i m \alpha\} \int_0^\infty R_m^*(r + r') f_m(r') r' dr'}{\iint_A f(r', \theta') r' dr' d\theta'} \quad (4.10)$$

Note that the squared modulus of the numerator in Eq. (4.10) attains its largest value when the input image is the same as the reference image, just as in Eq. (4.5), when the correlation signal is properly normalized. Also note that the mean value of the complex correlation signal is proportional to the high-light-level correlation signal. Because the rotation angle again appears only as a phase term, the squared modulus of $\langle C(r, \alpha) \rangle$ is invariant with respect to the orientation of the input image.

Even though the squared modulus of the mean value of the photon-limited correlation signal is rotation invariant, it is desirable to obtain a rotation-invariant estimate for the correlation signal with a single realization of the correlation signal, rather than perform enough realizations to obtain an accurate estimate for the mean value. Hence, the probability density function for the squared modulus of the correlation signal is required to predict probabilities of detection and false alarm for a given number of detected photoevents N . To find the probability density function for $|C|^2$, the following change of variables is made in Eq. (2.40): let

$$C'(x, y) = |C(x, y)| \cos \gamma, \quad C''(x, y) = |C(x, y)| \sin \gamma, \quad (4.11)$$

where $|C|^2$ and γ are given by

$$|C(x, y)|^2 = C'^2(x, y) + C''^2(x, y), \quad \gamma = \tan^{-1} \left(\frac{C'(x, y)}{C''(x, y)} \right). \quad (4.12)$$

The marginal density function for the squared modulus of the correlation signal $P(|C|^2)$ can then be written as

$$\begin{aligned}
 P(C^2) = & \frac{1}{4\pi(1-\rho^2)^{\frac{1}{2}}\sigma'\sigma''} \int_0^{2\pi} d\gamma \left\{ \right. \\
 & \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{|C|^2 \cos^2(\gamma) - 2|C| \langle C' \rangle \cos(\gamma) + \langle C' \rangle^2}{\sigma'^2} \right. \right. \\
 & - 2\rho \frac{|C|^2 \cos(\gamma) \sin(\gamma) - |C| [\langle C'' \rangle \cos(\gamma) + \langle C' \rangle \sin(\gamma)] + \langle C' \rangle \langle C'' \rangle}{\sigma' \sigma''} \\
 & \left. \left. + \frac{|C|^2 \sin^2(\gamma) - 2 \langle C'' \rangle |C| \sin(\gamma) + \langle C'' \rangle^2}{\sigma''^2} \right] \right\} \left. \right\} \quad (4.13)
 \end{aligned}$$

Unfortunately, a closed form solution to the integration in Eq. (4.13) does not exist in general²⁸. However, the integration is readily performed numerically. Equation (4.13) can be used to generate theoretical predictions for the probability distributions for the squared modulus of the correlation signal when the various objects are correlated with a reference circular harmonic.

If the input image is the same as the reference image, then one can compute the probability density function for the rotation angle α by noting that $\gamma = m\alpha$ [see Eqs. (4.8) and (4.12)]. The marginal density function for the rotation angle of the input is obtained using Eqs. (4.11), (4.12) and (2.40):

$$\begin{aligned}
 P(\alpha) = & \frac{1}{4\pi(1-\rho^2)^{\frac{1}{2}}\sigma'\sigma''} \int_0^{2\pi} d\phi \left\{ \right. \\
 & \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{|C|^2 \cos^2(m\alpha) - 2|C| \langle C' \rangle \cos(m\alpha) + \langle C' \rangle^2}{\sigma'^2} \right. \right. \\
 & - 2\rho \frac{|C|^2 \cos(m\alpha) \sin(m\alpha) - |C| [\langle C'' \rangle \cos(m\alpha) + \langle C' \rangle \sin(m\alpha)] + \langle C' \rangle \langle C'' \rangle}{\sigma' \sigma''} \\
 & \left. \left. + \frac{|C|^2 \sin^2(m\alpha) - 2 \langle C'' \rangle |C| \sin(m\alpha) + \langle C'' \rangle^2}{\sigma''^2} \right] \right\} \left. \right\} \quad (4.14)
 \end{aligned}$$

By using Eq. (4.14) it is possible to predict the accuracy in the determination of the rotation angle α for a given number of detected photoevents N .

4.4 Experimental Results

Experiments were performed to test the low-light-level performance of the complex circular-harmonic filter for both image discrimination and the determination of the rotation of the reference object. This was performed for the case of recognition of a single object in a dark background, and for the case of an object in a cluttered background. The system configuration is repeated in Fig. (4.1). In the first experiment, 35-mm format input scenes illuminated by an incoherent light source were imaged onto a two-dimensional, photon-counting detector. A neutral density of 10 was inserted between the input scene and detector to obtain an acceptable count rate; in this case the count rate was approximately 30 KHz. A dark count of approximately 50 Hz. was observed at room temperature.

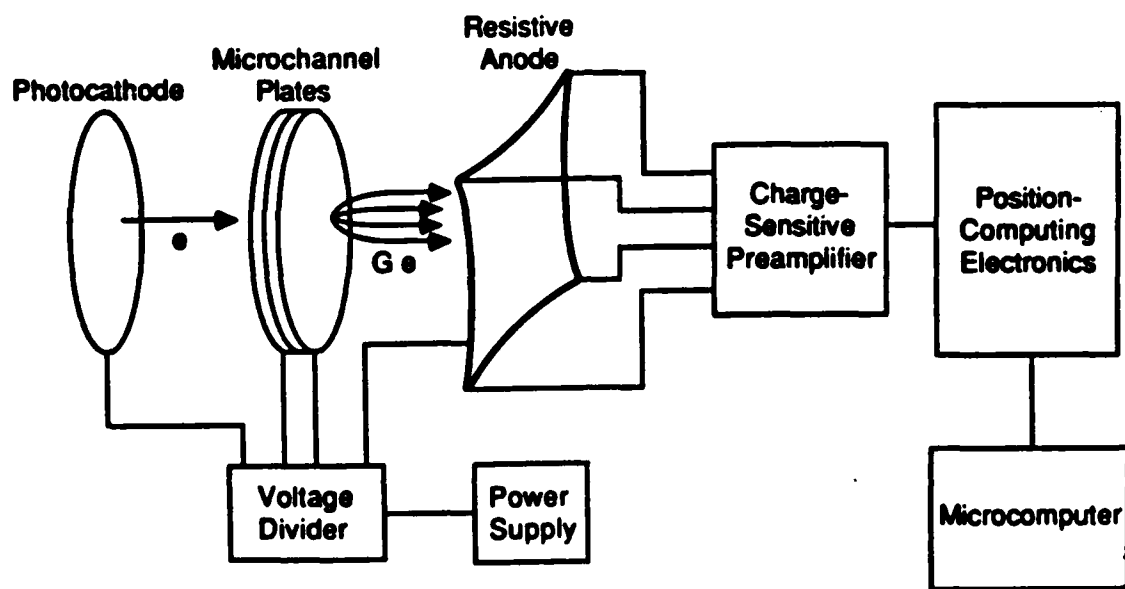


Figure 4.1 Schematic diagram of resistive-anode based photon-counting detection system.

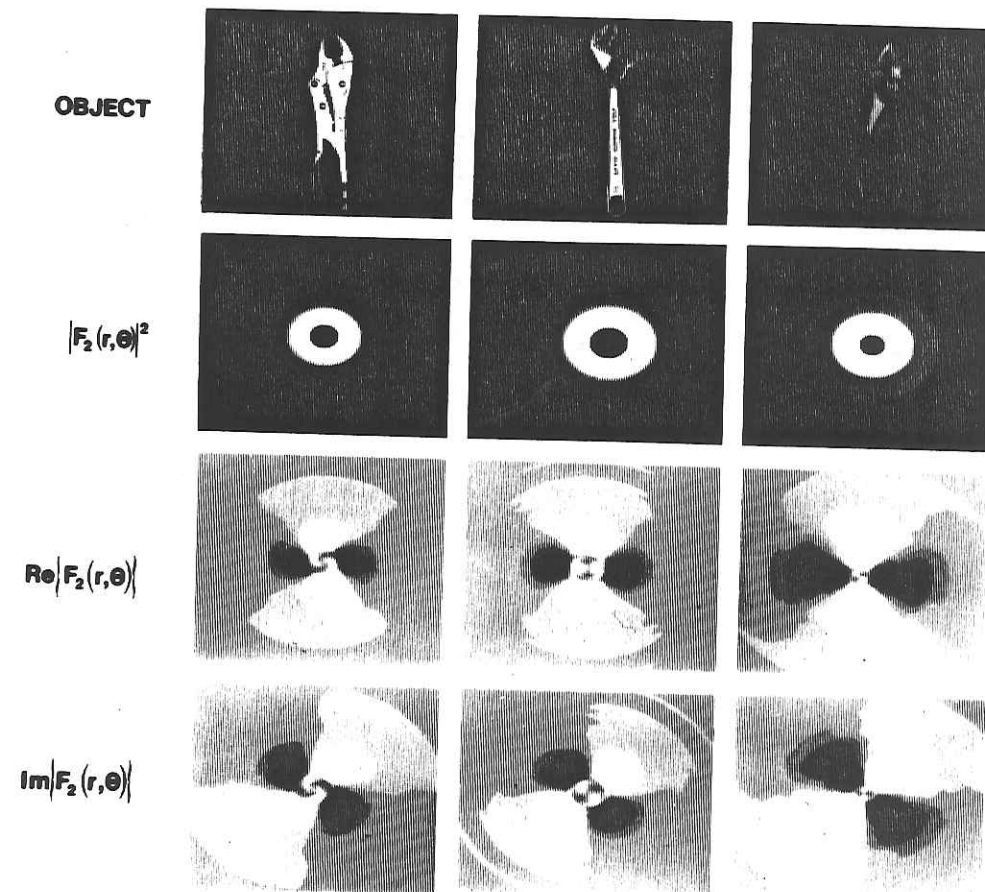


Fig. 4.2 Input Objects and associated second-order circular-harmonic components.

Figure 4.2 shows the input images and their associated circular-harmonic components used in the correlation experiments. The vise-grips were taken as the reference object, with a pliers and crescent wrench as test objects. In all of the experiments performed, the complex conjugate of the second circular-harmonic of the vise-grips (computed about the centroid of the object) was chosen as the reference function (see Eq. (4.1) and Fig. 4.2). The correlation signals were realized by detecting a fixed number of photoevents, as discussed in Chapter 2. Equation (2.9) was used to compute the correlation signal and the rotation angle of the input with respect to the reference was estimated using Eq. (4.8).

To demonstrate the recognition performance of the circular-harmonic filter at low light levels, the complex correlation signal was computed with the input image rotated through angles of 0, 90, 180, and 270 degrees with respect to the reference image. One thousand measurements of the correlation signal were performed for each input scene (at each orientation) for different values of N ($N=500, 1000, 2000$ and 3000 detected photoevents).

The mean values and standard deviations of the probability density function for the squared modulus of the correlation signal obtained by using each input scene for $N=3000$ detected photoevents are shown in Table 4.1. Note that the mean value of the squared modulus of the correlation signal is the same for each orientation of the vise-grips.

Input Image		Theory		Experiment	
$f(x',y')$		$\langle C ^2 \rangle$	σ	$\langle C ^2 \rangle$	σ
Vise Grips:	$\alpha=0^\circ$	1.60E9	7.85E7	1.51E9	7.47E7
Vise Grips:	$\alpha=90^\circ$	1.60E9	7.85E7	1.58E9	7.67E7
Vise Grips:	$\alpha=180^\circ$	1.60E9	7.85E7	1.58E9	7.68E7
Vise Grips:	$\alpha=270^\circ$	1.60E9	7.85E7	1.58E9	7.55E7
Wrench:	$\alpha=0^\circ$	9.71E8	6.07E7	9.69E8	6.39E7
Pliers:	$\alpha=0^\circ$	7.06E8	5.36E7	7.03E8	4.55E7

Table 4.1. Comparison of theoretical predictions and experimental results for the expected values and standard deviations of the squared modulus of the correlation signal. The number of detected photoevents is $N=3000$.

A histogram of the correlation values obtained using 3000 detected photoevents is shown in Fig. 4.3. In the figure, I indicates the range of correlation values observed with the vise-grips input rotated by 90 degrees with respect to the reference, II shows the range of correlation values with the wrench input, and III shows the range of correlation values obtained with the pliers input. The solid curves represent theoretical predictions for the probability density function for the squared modulus of the correlation signal $|C|^2 = |C(0,0)|^2$. These curves are obtained by performing the integration in Eq. (4.13) numerically, using expressions for the mean values and variances given in Eqs. (2.41)-(2.42) and (2.48)-(2.50).

Figure 4.4 shows a histogram of values of the orientation angle α obtained from 1000 measurements of the complex correlation signal realized by using 3000 detected photoevents with the vise-grips rotated by 90 degrees with respect to the reference. The rotation angle α is obtained from the correlation signal using Eq. (4.8). The solid curve is the theoretical prediction of the probability density function for the rotation angle α . The solid curve is obtained using Eq. (4.14), with Eqs (2.41)-(2.42), and (2.48)-(2.50). In the experiment, the mean of $P(\alpha)$ was 89.6 degrees, and the standard deviation was 0.705 degrees.

In Fig. 4.5, the probability of false alarm and the probability of detection (ROC curves) are plotted for the vise grips (input rotated by 90 degrees) and for the crescent wrench for different values of N . Note that as the number of detected photoevents used to realize the correlation signal is increased, the recognition performance increases. For $N = 3000$, it is possible to set a decision threshold such that the probability of making an incorrect decision is less than 1×10^{-5} .

With a fixed number of detected photoevents, the standard deviation of the real and imaginary parts of the correlation signal (see Eqs (2.48)-(2.49)) is less than for the case when N is Poisson-distributed. This characteristic is true for the modulus of the correlation signal as well. This is demonstrated in Fig. 4.6. The two density functions in Fig. 4.6 are theoretical predictions for the squared modulus of the correlation when the vise grips are input at an angle of 90 degrees with respect to the reference. The smaller standard deviation in the modulus of the correlation signal results in a larger separation of the density functions using smaller numbers of detected photoevents. Hence, if one has the choice of realizing the correlation signal using either a fixed number of detected photoevents or a Poisson-distributed number of photoevents, one would usually choose to have N fixed.

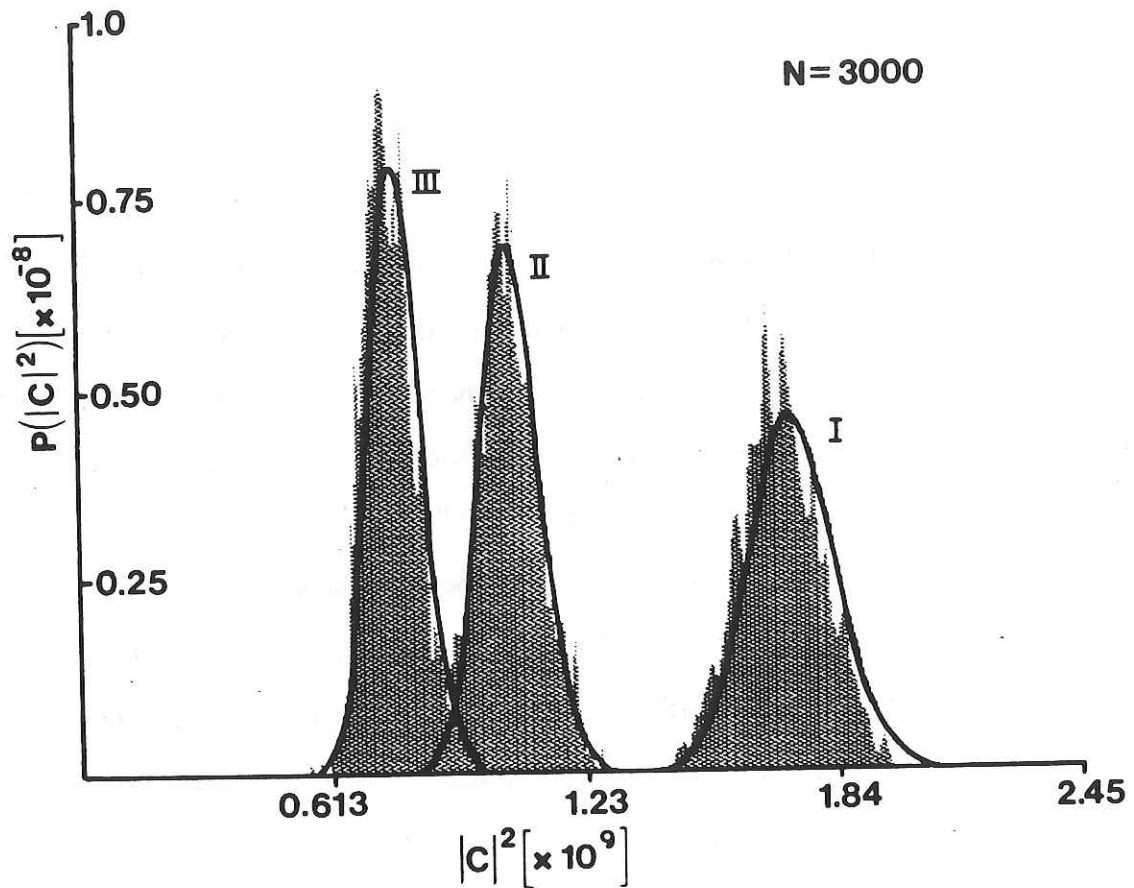


Figure 4.3 Histogram of experimental values of the squared modulus of the correlation signal when the input image is I, vise grips, rotated by 90 deg. with respect to the reference; II crescent wrench; and III, pliers. The reference function is the second circular-harmonic of the vise grips (see Fig. 4.2), and the number of detected photoevents is $N=3000$.

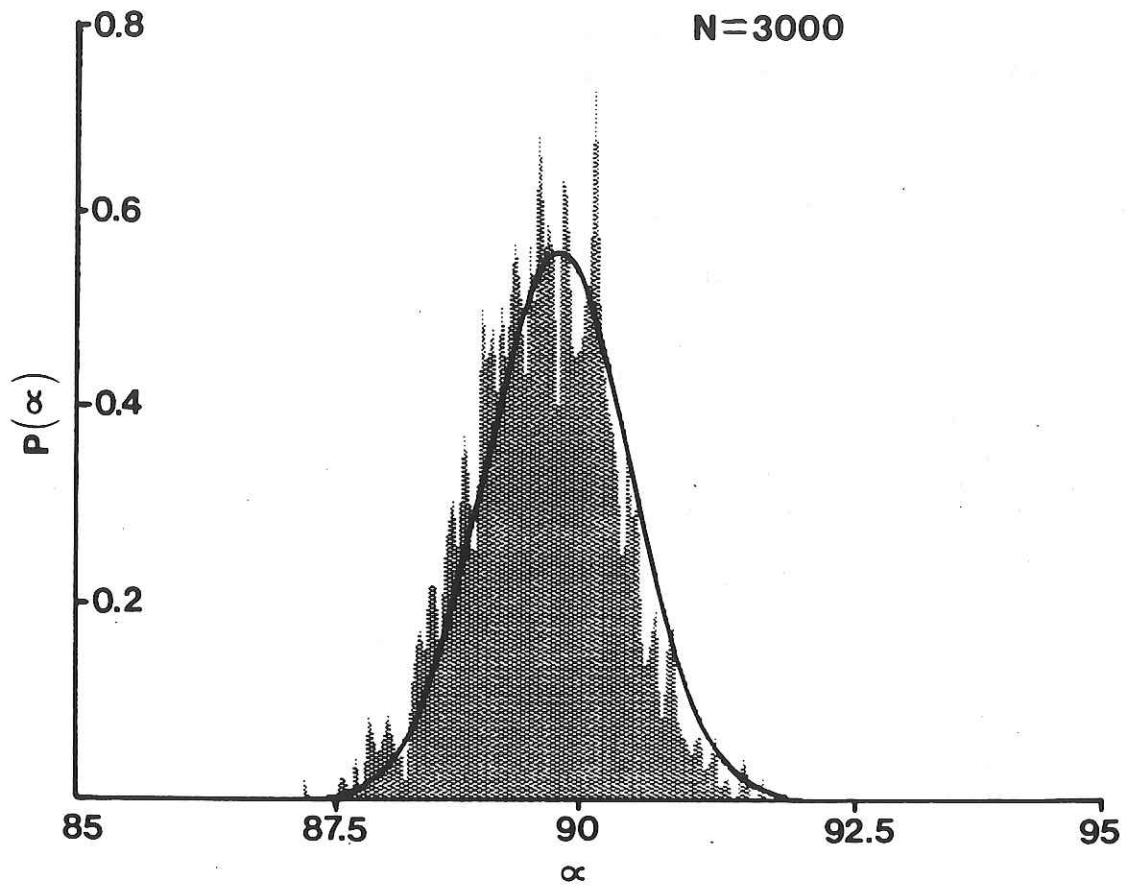


Figure 4.4 Histogram of values for the rotation angle α when the vise grips are input rotated by 90 degrees with respect to the reference. The number of detected photoevents is $N=3000$.

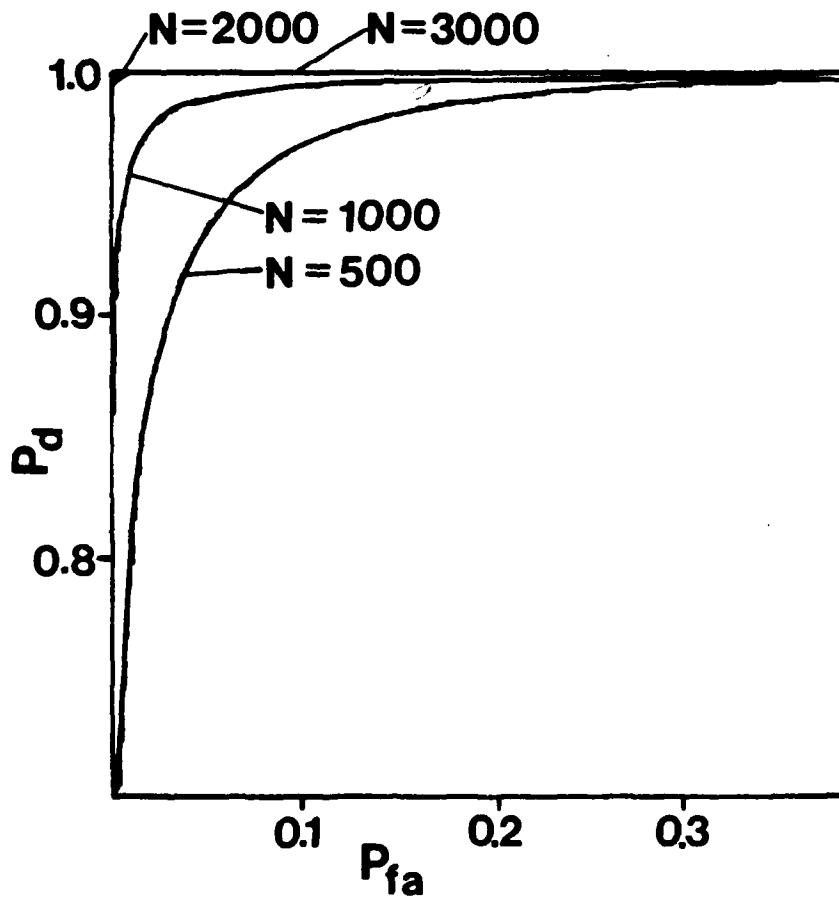


Figure 4.5 ROC curves for the vise grips and crescent wrench. The vise grips are rotated by 90 degrees with respect to the input.

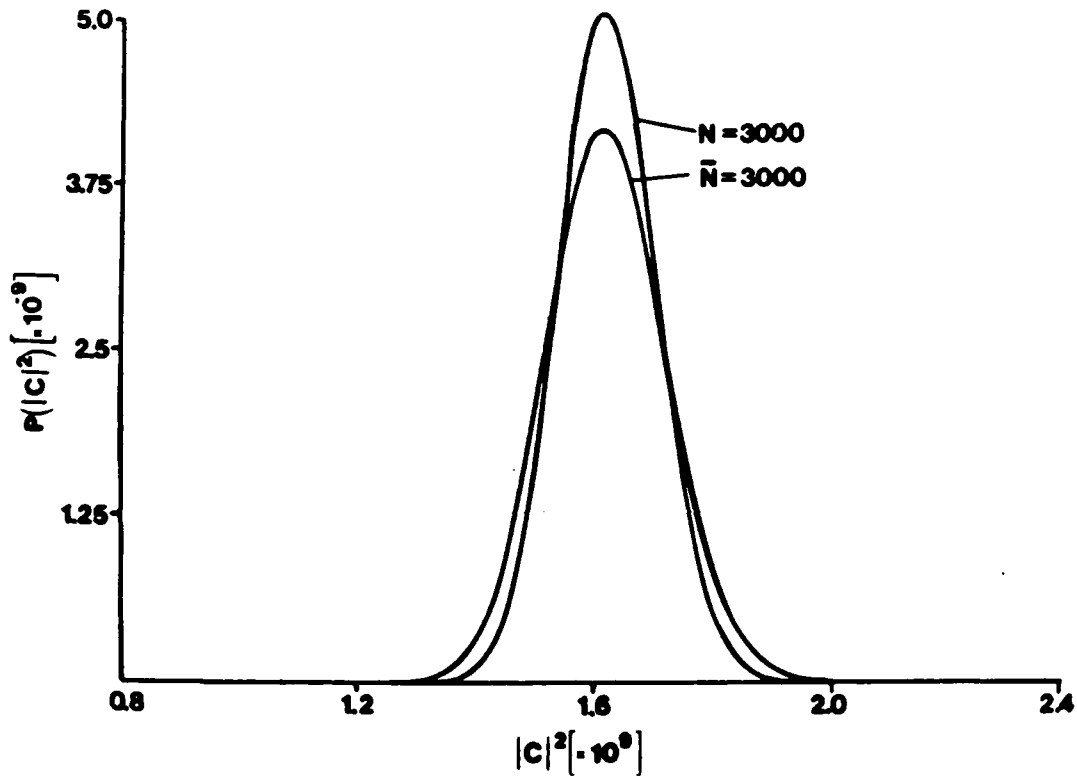


Figure 4.6 Comparison of theoretical probability density functions for the squared modulus of the correlation signal when the number of detected photoevents is fixed ($N=3000$) and when the number of detected photoevents is Poisson distributed ($\bar{N}=3000$). The input image is the vise grips rotated by 90 degrees with respect to the input.

4.5. Rotation-Invariant Recognition in a Cluttered Environment

4.5.1 Introduction

Past work^{6,10} using circular-harmonic filters in a cluttered environment has been largely unsuccessful. The primary reason for this fact is that in conventional implementations of the cross correlation, it is difficult to perform the proper normalization of the correlation signal. The proper normalization of the cross-correlation function between an input scene and the circular-harmonic filter is given by the Schwarz inequality. The Schwarz inequality²⁹, describes the normalization required for the cross-correlation between two functions $A(r, \theta)$ and $B(r, \theta)$ to attain a maximum at the origin; i.e.

$$\left| \int_0^{2\pi} \int_0^{\infty} A(r+r', \theta+\theta') B^*(r', \theta') r' dr' d\theta' \right|^2 \leq \left(\int_0^{2\pi} \int_0^{\infty} |A(r', \theta')|^2 r' dr' d\theta' \right) \left(\int_0^{2\pi} \int_0^{\infty} |B(r', \theta')|^2 r' dr' d\theta' \right) \quad (4.15)$$

Note that relation (4.15) attains a maximum when A is equal to βB , where β is a constant. Examining Eq. (4.5), one sees immediately that the proper normalization for the squared modulus of the correlation function is given by

$$|C(r, \alpha)|^2 \equiv \frac{\left| e^{im\alpha} \int_0^{R_{max}} F_m^*(r+r') R_m(r') r' dr' \right|^2}{\left(\int_0^{R_{max}} |F_m(r')|^2 dr' \right) \left(\int_0^{R_{max}} |R_m(r')|^2 dr' \right)} \quad (4.16)$$

where $F_m(r')$ is the radial part of the m^{th} circular-harmonic of the input scene, computed about each location in the input scene. $R_m(r')$ is the radial part of the

circular-harmonic reference function, and is obviously the same for all points in the correlation function. In Eq. (4.16), the limits of integration have been truncated to a value R_{\max} , which is typically determined by the location of the proper expansion center of the reference object, and by the size of the reference object. Note that a different normalization is required for each point in the correlation function. For example, if the input scene is 512x512 pixels, and the reference circular-harmonic function is 64x64 pixels, then $G_m(r')$ must be computed [See Eq. (4.2)], and integrated at 28,672 points. Clearly, this is very difficult to perform in real-time with most digital processors, and is also difficult to perform optically.

In the low-light-level implementation, the mean value of the correlation signal was proportional to the numerator of Eq. (4.16), but was divided by the integrated intensity of the input scene that fell within the reference function window [see Eq. (4.10)]. This 'energy normalization' occurs naturally as a result of the photon statistics. While this is not the type of normalization that will guarantee that the correlation function will attain a maximum, it will prevent very bright areas of the input scene from causing the correlation signal to exceed a predetermined threshold and cause a false alarm. This energy normalization was sufficient to allow excellent recognition performance in the examples discussed in Section 4.4. However, this normalization is often insufficient to achieve accurate detection of objects from within a cluttered scene, such as that shown in Fig. 4.7. For example, in this case the modulus of the mean value of the correlation signal (Eq. (4.10)) did not allow the trucks to be recognized from within the clutter present in the scene. Using photon-counting techniques, it is possible to obtain a real-time normalization that will improve the recognition of the performance of the circular-harmonic filter when operating in cluttered environments, such as Fig. 4.7.

4.5.2 Normalization using Photon-Counting Techniques

Using photon-counting techniques, it is possible to compute a normalization for the output correlation signal that is superior to the "energy" normalization that occurs naturally as a result of the photon statistics. However, a method of computing the normalization suggested by the Schwarz inequality in real time has not been obtained. In this section, the method for computing this improved normalization is described.

As indicated in Section 4.5.1, normalization by the quantity suggested by the Schwarz inequality requires that the circular-harmonic of the input be computed at every point in the input scene. By choosing the proper reference function, it is possible to compute an approximation for the Schwarz inequality normalization.

If a photon-counting detection system is used to recognize multiple objects from within a cluttered environment, it is necessary to utilize a reference scene window. This is due to the fact that the correlation filtering is performed in an image plane; as a result, a search of the input scene must be performed by moving the reference window throughout the input scene. An effective means of performing this search is discussed in Chapter 6; for now, let us assume that the reference window is moved sequentially over every possible point in the input, and it desired to compute the normalization factor at each point.

Consider the case in which one chooses, in parallel, two different reference functions $R_1(r, \theta)$ and $R_2(r, \theta)$. Let R_1 be a circular-harmonic component of the reference function, computed about its proper center. In addition, let $R_2(r, \theta)$ be given by

$$R_2(r, \theta) = \exp - (im\theta) \quad , \quad (4.17)$$

where the origin of the polar coordinate system is taken to be the location in the input scene for which the normalization factor is being calculated. When the photon-limited correlation signal is calculated at a given offset, a fixed number of detected photoevents N are detected in Cartesian coordinates, and then converted to polar coordinates relative to the location of the reference window within the input scene. The correlation signals $C_1(0,0)$ and $C_2(0,0)$ are calculated using Eq. (2.9), with the reference functions given in Eqs. (4.3) and (4.16), respectively. The mean value of C_1 is once again given by Eq. (4.10). The mean value of C_2 is given by [using Eqs. (2.43) and (2.44)]

$$\langle C_2(0,0) \rangle = \frac{\int_0^{R_{\max}} \int_0^{2\pi} f(r', \theta') e^{-im\theta'} r' dr' d\theta'}{\int_0^{R_{\max}} \int_0^{2\pi} f(r', \theta') r' dr' d\theta'} \quad (4.18)$$

Using Eq. (4.2), Eq. (4.18) can be re-written as

$$\langle C_2(0,0) \rangle = \frac{\int_0^{R_{\max}} F_m(r') r' dr'}{\int_0^{R_{\max}} \int_0^{2\pi} f(r', \theta') r' dr' d\theta'} \quad (4.19)$$

Hence, dividing the squared modulus of $\langle C_1 \rangle$ by the squared modulus of $\langle C_2 \rangle$, and dividing by the integral of the squared modulus of $R_m(r)$, one obtains

$$\frac{|\langle C_1 \rangle|^2}{|\langle C_2 \rangle|^2} = \frac{\left| \int_0^{R_{\max}} F_m(r') R_m(r') r' dr' \right|^2}{\left| \int_0^{R_{\max}} F_m(r') r' dr' \right|^2 \int_0^{R_{\max}} |R_m(r')|^2 r' dr'} \quad (4.20)$$

Comparing Eq. (4.20) with Eq. (4.16), one sees immediately that the normalization obtained using the photon-counting technique is slightly different than that required by

the Schwarz inequality. In the photon-limited case, the normalization factor in the denominator is

$$\left| \int_0^{R_{\max}} F_m(r') r' dr' \right|^2, \quad (4.21)$$

while the Schwarz inequality requires that

$$\int_0^{R_{\max}} |F_m(r')|^2 r' dr' \quad (4.22)$$

be present in the denominator. Because the Schwarz inequality is not satisfied with the photon-limited normalization, it is not guaranteed that the peak value(s) of the correlation function for an entire input scene will be the result of the cross-correlation with the reference object(s). However, in the case of the cluttered input scene shown in Fig. 4.7, encouraging results were obtained using this normalization; with the original energy normalization, accurate identification of the trucks in Fig. 4.7 was not possible. As a result, the effectiveness of this normalization must be tested on a case-by-case basis.

Just as in Section 4.4, it is desirable to obtain an estimate for the normalized correlation signal using a single realization of the correlation signal, rather than perform a large number of realizations and obtain the mean values indicated in Eq. (4.20). At a given position in the input scene, the estimate for the normalized correlation signal will be obtained by detecting a fixed number of photoevents, and computing correlations $C_1(0,0)$ $C_2(0,0)$ using Eq. (2.9), with the reference functions given in Eqs. (4.3) and (4.16), respectively. One then computes the ratio $|C_1(0,0)|^2 / |C_2(0,0)|^2$ to determine the estimate for the normalized correlation signal. Hence, it is necessary to obtain an expression for the probability density function for the ratio of the moduli of the two

correlation signals, if one is to predict the recognition performance for a given reference function and input scene.

We define the new variable Z such that

$$Z \equiv \frac{\xi}{\eta}; \quad \xi \equiv |C_1|^2, \quad \eta \equiv |C_2|^2. \quad (4.23)$$

If different photoevents are used to realize each correlation signal, then the random variables ξ and η will be statistically independent. Hence, the density function for $P(Z)$ will be given by³⁰

$$P(Z) = \int_0^{\infty} \eta P_{\xi}(Z\eta) P_{\eta}(\eta) d\eta, \quad (4.24)$$

where $P_{\xi}(Z\eta)$ is obtained by substituting $Z\eta$ for $|C|^2$ in Eq. (4.13). The expressions for the mean values and variances are obtained using Eqs. (2.43), (2.44), (2.48), (2.49) and (2.50), with the reference function given in Eq. (4.3). Similarly, the probability density function $P_{\eta}(\eta)$ is obtained by substituting η for $|C|^2$ in Eq. (4.13). In this case, the mean values and variances are obtained using the reference function given in Eq. (4.17). Clearly, the integrations in Eq. (4.24) cannot be performed analytically, however it is straightforward to perform them numerically. It is important to note that these calculations are performed in advance of any recognition experiments, and is not a limitation to this method for image recognition.

A typical environment that was tested is illustrated in Fig. 4.8. The input scene (lower left) (512x512 pixels) is a portion of the aerial photograph shown in Fig. (4.7). The reference object (64x32 pixels) (top left) was chosen to be the top truck in the center of the input scene. The reference function used was the 2nd circular-harmonic of the truck, computed about its proper center. The proper center was chosen using the

technique described in Section 4.2. Theoretical predictions for the squared modulus of the mean value of the normalized output correlation signal divided by the number of detected photoevents is shown in the lower right of Figure 4.7. This is analogous to the high-light-level correlation output. The mean values at each point in the correlation function were obtained using Eq. (4.24). Note that, properly thresholded, the output correlation signal can be used to identify four of the six trucks in the input scene. Only the "double truck" located in the center of the input scene is not identified.

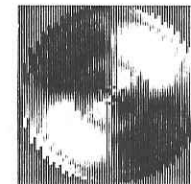
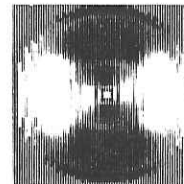
Reference Image



64 x 28 (magnified)

Filter [64 x 64 (magnified)]

Second Circular Harmonic



Real

Imaginary

Input (512 x 512)



Correlation Output

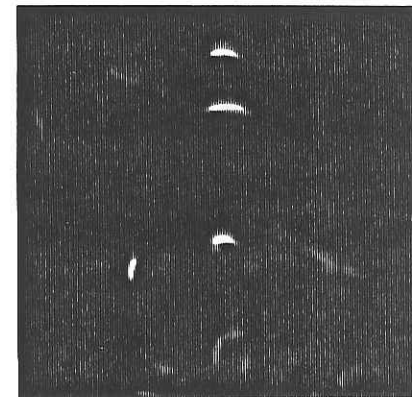


Figure 4.7 Rotation-invariant filtering in a cluttered environment. Top left, reference object; top right, second circular-harmonic of reference object; bottom left, input scene; bottom right, output correlation function.

4.6. Discussion

The experimental results obtained show that rotation-invariant image recognition can be achieved with quantum-limited images. Excellent experimental agreement was observed between the theoretical predictions for both the probability density functions of the output correlation signal and orientation angle α .

One advantage in the use of circular-harmonic components of a reference image for rotation-invariant image recognition is that the rotation angle of the reference object can be obtained from the phase of the correlation signal. However, depending on the order of the harmonic used as the reference, there are certain ambiguities in the rotation angles that are determined from the correlation signal. For example, let $m=2$ in Eq. (4.6), and assume that the reference object is input, which makes $C(r,0)$ real. Equation (4.6) can then be rewritten as

$$C(r,\alpha) = C(r,0)[\cos(2\alpha) + i \sin(2\alpha)] \quad (4.25)$$

Note that the phase of the correlation signal will have the same value if $\alpha'=\alpha+\pi$. Hence, the orientation of the reference is not uniquely determined. (This result can also be obtained by observing the 180 degree symmetry in the real and imaginary parts of the second harmonic components in Fig. 4.1.) When the correlations are performed at low light levels, this problem is easily solved. One simply rotates the coordinates of a few detected photoevents by $\pi/4$ radians, and observes the change (if any) in the sign of the phase of the correlation signal. Using this sign information the rotation angle can be uniquely determined. It should be stressed that only a few tens of detected photoevents need be rotated, and that the angle through which they are rotated depends on the order of the harmonic that is used as the reference.

4.7 Summary

The recognition performance of the circular-harmonic filter is investigated at low light levels. Theoretical expressions are given in Section 4.3 for the probability density function of the correlation signal realized by cross-correlating a photon-limited input image with a circular-harmonic of a reference image. The theoretical predictions are verified experimentally in Section 4.4 (see Figs. 4.3 and 4.4), with excellent agreement between theory and experiment. In Section 4.5, the circular-harmonic filter is employed in an application to recognition within a cluttered environment, with encouraging results (see Fig. 4.8).

References for Chapter 4

1. A. B. Van der Lugt, 'Signal detection by complex spatial filtering,' IEEE Trans. IT-10, 139-145 (1964).
2. D. Casasent and D. Psaltis, 'Position, rotation and scale invariant optical correlation,' Appl. Opt. **15**, 1795-1799 (1976).
3. Y. Hsu and H. H. Arsenault, 'Optical pattern recognition using circular-harmonic expansion,' Appl. Opt. **21**, 4016-4019 (1982).
4. R. Wu and H. Stark, 'Rotation-invariant pattern recognition using optimum feature extraction,' Appl. Opt. **24**, 179-184 (1985).
5. K. Mersereau and G. M. Morris, 'Scale, rotation and shift-invariant image recognition,' Appl. Opt. **21**, 2338-2342 (1986).
6. H.H. Arsenault and Yunlong Sheng, 'Properties of the circular-harmonic expansion for rotation-invariant pattern recognition,' Appl. Opt. **25**, 3225-3229 (1986).
7. Yunlong Sheng and H.H. Arsenault, 'Method for determining expansion centers and predicting sidelobe levels for circular-harmonic filter,' J. Opt. Soc. Am. A **4**, 1793-1797 (1987).
8. M.T. Manry, 'Compound filter using circular harmonic expansion,' Appl. Opt. **26**, 3622-3627 (1987).
9. G. F. Schils and D. W. Sweeney, 'Rotationally-invariant correlation filtering,' J. Opt. Soc. Am. A **2**, 1411-1418 (1985).
10. G. F. Schils and D. W. Sweeny, 'Rotationally-invariant correlation filtering for multiple objects,' J. Opt. Soc. Am. A **3**, 902-909 (1986).
11. M.T. Manry, 'Compound filter using circular harmonic expansion,' Appl. Opt. **26**, 3622-3627 (1987).
12. A. Metioui, H.H. Arsenault and L. Leclerc, 'Methods for reducing sidelobes associated with composite filters,' Opt. Commun. **71**, 332-336 (1989).
13. H.H. Arsenault and Yunlong Sheng, 'Modified composite filter for pattern recognition in the presence of noise with a non-zero mean,' Opt. Commun. **63**, 15-20 (1987).
14. L. Leclerc, Yunlong Sheng and H.H. Arsenault, 'Rotation-invariant phase-only and binary-phase-only correlation,' Appl. Opt. **28**, 1251-1256 (1989).

15. H. J. Caulfield, 'Linear combinations of filters for character recognition: a unified treatment,' Appl. Opt. **19**, 3877-3878 (1980).
16. B.V.K. Kumar, 'Efficient approach to designing linear combination filters,' Appl. Opt. **22**, 1445-1448 (1983).
17. D. Casasent, 'Unified synthetic discriminant function,' Appl. Opt. **23**, 1620-1627 (1984).
18. B.V.K. Kumar and E. Pochapsky, 'Signal-to-noise ratio considerations in modified matched spatial filters,' J. Opt. Soc. Am. A **3**, 777-786 (1986).
19. B.V.K. Kumar, 'Minimum-variance synthetic discriminant function,' J. Opt. Soc. Am. A **3**, 1579-1584 (1986).
20. R. J. Marks II, J.A. Ritcey, L. E. Atlas and K. F. Cheung, 'Composite matched filter output partitioning,' Appl. Opt. **26**, 2274-2278 (1987).
21. D. Casasent, 'Correlation synthetic discriminant functions,' Appl. Opt. **25**, 2343-2350 (1986).
22. J. L. Horner and P. D. Gianino, 'Phase-only matched filtering,' Appl. Opt. **24**, 2889-2893 (1985).
23. L. Horner and J. Leger, 'Pattern recognition with binary phase-only filters,' Appl. Opt. **23**, 812-816 (1984).
24. J. L. Horner and P. D. Gianino, 'Applying the phase-only filter to the synthetic discriminant function correlation filter,' Appl. Opt. **24**, 851-855 (1985).
25. E. Paek and D. Psaltis, 'Optical associative memory using Fourier transform holograms,' Opt. Eng. **26**, 428-433 (1987).
26. L. Saaf and G. M. Morris, 'Back-propagation filters for image classification,' submitted to Appl. Opt., Nov. (1989).
27. T.A. Isberg and G.M. Morris, 'Rotation-invariant image recognition at low light levels,' J. Opt. Soc. Am. A **3**, 954-963 (1986).
28. K. S. Miller, *Multivariate Distributions* (Krieger, New York, 1975) p. 28.
29. G. Arfken, *Mathematical Methods for Physicists*, (Academic, New York, 1970) p.445.
30. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1965), p.197.

Chapter 5

Monte Carlo Estimation of Moment Invariants for Pattern Recognition

5.1 Introduction

Moment invariants have been used for invariant pattern recognition for many years. Since introduced by Hu¹, moment techniques for pattern recognition have been applied to image recognition problems ranging from scene matching in satellite images² to character and ship recognition³⁻⁶, with varying degrees of success. Recent derivations of complex moment invariants^{7,8} have yielded improved results and have renewed interest in this method for image recognition.

One major drawback to the use of moment invariants for pattern recognition is the massive amount of numerical computations necessary to determine the invariant moments for a given image; this computational complexity makes it difficult to compute moment invariants in real time². Several authors have demonstrated novel optical-digital hybrid methods for computing invariant moments¹⁰⁻¹⁵. These methods often employ the use of coherent scene illumination, which may be impractical in some applications. In this chapter, a real-time Monte Carlo method for estimation of moment invariants that works with incoherent input scene illumination is presented. It is demonstrated that estimates for moments of circular harmonic functions¹⁶⁻¹⁸ can be determined in real time using a position-sensitive photon-counting detection system¹⁹ in which the locations of individual detected photoevents serve as the source of random numbers for the Monte Carlo algorithm²⁰⁻²¹. The detection system is intensity based, which eliminates the need for coherent illumination; virtually any

incoherent source is sufficient. This method for moment computation has the added advantage that the moment invariants are computed without digitizing the input scene.

Herein, we demonstrate that the information provided by the spatial coordinates of a few thousand detected photoevents can yield an accurate estimate for complex moment invariants of a given input image. The detection system employed here can detect photoevents at rates up to 100 kHz., which makes it possible to compute estimates for complex moment invariants in a few tens of milliseconds. It is the speed of the photon-counting detection system coupled with the fact that only a small number of detected photoevents are necessary to get an accurate estimate for the moment invariants that makes this photon-counting technique advantageous. Theoretical predictions for the accuracy to which moment invariants can be determined for a given number of detected photoevents are described, and tested experimentally.

The basic technique of moment invariants for pattern recognition is briefly reviewed in Section 5.2. In Section 5.3, we present the results of the theory for computing estimates for moment invariants using photon-counting techniques; the details of the relevant statistics are provided in Appendices A and B. Experimental confirmation of the theory is given in Section 5.4.

5.2 Moment Invariants for Pattern Recognition

5.2.1 Basic Definitions

The use of invariant moments for pattern recognition was first introduced by Hu¹. The invariant moments are based on the geometrical moments m_{pq} of an image,

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) x^p y^q dx dy \quad , \quad (5.1)$$

where $f(x,y)$ is the input image in question and p and q are non-negative integers. Hu defines a set of moment invariants that are a nonlinear combination of the geometrical moments; these moment invariants are theoretically invariant with respect to shift, scale and rotation. The major limitations of Hu's moments are the large dynamic range of the various orders², the accuracy to which the moments must be computed to produce adequate recognition capabilities¹⁵ and their poor performance under less than optimum (i.e. noisy) conditions². It has been demonstrated that accurate computation of Hu's moment invariants requires values for the geometrical moments to be accurate within 1%^{2,9,15}. This accuracy requirement makes the computation of the geometrical moments particularly susceptible to quantization errors in digital computations^{2,9}, and is the limiting factor in the design of optical-digital hybrid systems for computing Hu's moment invariants.

Several variations on Hu's moments have been suggested. Reddi²⁴ suggested a change to polar coordinates, which allows radial and angular moments to be computed. This type of moment is less susceptible to quantization errors²⁴. More recently, Abu-Mostafa and Psaltis^{7,8} proposed the use of complex geometrical moments C_{pq} , defined as

$$C_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) (x+iy)^p (x-iy)^q dx dy \quad . \quad (5.2)$$

They showed that these complex moments could be used to define a set of moment invariants, and related the complex moments to certain circular-harmonic components¹⁶ of the image.

5.2.2 Radial Moments of Circular-Harmonic Functions (CHF's)

Sheng and Arsenault¹⁷ and Sheng and Duvernoy¹⁸ showed that a particular case of a general descriptor, namely a Fourier-Mellin descriptor, produced a more general complex moment invariant; i.e., they defined a moment invariant by using radial moments of the circular-harmonic components of an image. This moment invariant combines many of the desirable properties of the radial and angular moments described by Reddi²⁴, and the complex moments defined by Abu-Mostafa and Psaltis^{7,8}.

The Fourier-Mellin descriptors $M_{s,m}$ are defined as

$$M_{s,m} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} r^{s-1} f(r,\theta) \exp(-im\theta) dr d\theta \quad , \quad (5.3)$$

where m is an integer, and s is, in general, complex. Note that this is actually a radial Mellin transform of the m^{th} circular harmonic component of the input scene $f(r,\theta)$. However, in the special case in which s is an integer, $M_{s,m}$ becomes a radial moment of the circular-harmonic components of the image. The descriptors $M_{s,m}$ are used to define moment invariants $\Phi_{s,m}$ as follows:

$$\Phi_{s,m} = \frac{|M_{s,m}|}{|M_{s,0}|} \quad . \quad (5.4)$$

The moments $\Phi_{s,m}$ are invariant with respect to rotation and scale, and are invariant with respect to position when the $M_{s,m}$ are computed about the centroid of the input image. Taking the modulus of $M_{s,m}$ provides rotation invariance, while the normalization by the 0th order radial moment provides scale invariance. For example, consider the case in which the input image $f(r,\theta)$ is scaled by a factor α , rotated through an angle β , and multiplied by a contrast factor κ . The moment of the modified image $M'_{s,m}$ becomes

$$M'_{s,m} = \frac{\kappa}{2\pi} \int_0^{2\pi} \int_0^\infty r^{s-1} f(\alpha r, \theta + \beta) \exp(-im\theta) dr d\theta \quad (5.5)$$

If one multiplies and divides Eq. (5) by $[\alpha^{-s+1} \exp(im\beta)]$, one obtains

$$M'_{s,m} = \frac{\kappa}{2\pi} \exp(im\beta) \alpha^{-s} \left\{ \int_0^{2\pi} \int_0^\infty (\alpha r)^{s-1} f(\alpha r, \theta + \beta) \exp[-im(\theta + \beta)] d(\alpha r) d(\theta + \beta) \right\} \quad (5.6).$$

If one makes the following change of variables in Eq. (5.6): $r' = \alpha r$, $\theta' = \theta + \beta$, and compares the result with Eq. (5.3) one obtains

$$M'_{s,m} = \alpha^{-s} \exp(im\beta) M_{s,m} \quad (5.7)$$

In a similar manner, one obtains for the 0th order radial moment $M'_{s,0}$

$$M'_{s,0} = \alpha^{-s} \kappa M_{s,0} \quad (5.8)$$

Hence, using Eqs. (5.4), (5.7) and (5.8), one sees immediately that $\Phi'_{s,m} = \Phi_{s,m}$ and is therefore invariant with respect to the above geometric distortions.

The moment invariants are taken as features of the image to be recognized, and a recognition decision is based on the distance in feature space between the features of an input image and a reference image. This distance D^2 is defined as

$$D^2 = \sum_m [\Phi_{2,m}^{\text{input}} - \Phi_{2,m}^{\text{ref}}]^2, \quad (5.9)$$

where Φ^{input} and Φ^{ref} denote the features for the input and reference images, respectively. A recognition decision is made by setting some threshold distance D_T in feature space. If the observed value for D^2 is greater than D_T , the input image is said to be different than the reference image; if D^2 is less than D_T , the observed image is said to be the same as the reference image.

As mentioned earlier, when the variable s is taken to be an integer, the Fourier Mellin descriptors are really complex moment invariants, in which one computes radial moments of circular-harmonic components of the input image. The performance of this particular type of moment invariant is superior to the performance of complex moment invariants for the following reason: classical moment invariants¹⁻⁶ and complex moments^{7,8}, which are defined in Cartesian coordinates, can be shown¹⁷ to be special cases of radial moments of CHF's. For example, the complex moments defined in Ref. 7 are equal to $M_{s,m}$ with $s=|m|+2$, $|m|+4$, $|m|+6$ In this case, s and m are constrained to have particular values, which may limit their performance.

Low-order radial moments have been shown to be superior¹⁷ to higher order moments for pattern recognition, particularly in realistic scenes where noise may be present. For a given set of objects, Sheng and Arsenault¹⁷ reported a 30% misclassification rate using moments with $s=3$; no mistakes were made when moments with s equal to 1 or 2 were used. In addition, because the radial moments do not form an orthogonal set, radial moments of different order contain much of the same information about the input image. Because of this information redundancy, many times one does not gain new information about the object in question by using higher-order radial moments. Hence, the best strategy is often to choose the low-order radial

moments and choose multiple circular-harmonic orders¹⁷. Using these low-order radial moments, it is often possible to achieve better over-all recognition performance than is possible with complex moments that are defined in Cartesian coordinates¹⁷. For this reason, we will only consider radial moments of CHF's in this paper.

The difficulty in real-time digital computation of moment invariants has led a number of researchers to investigate the possibility of using an opto-electronic hybrid system to compute moment-invariants in real time. Casasent and Psaltis¹⁰ described a hybrid processor in which the input scene was coherently illuminated and spatially convolved with a single fixed mask. This allowed the moments to be computed in parallel. Teague¹³ described a system that used coherent optical preprocessing to compute moments of the input scene. Wagner and Psaltis¹² suggested a system using acousto-optics to compute image moments, and Kumar and Rahenkamp¹¹ demonstrated a system that computed geometric moments using Fourier-plane intensities. Finally, Duvernoy and Sheng¹⁵ proposed an incoherent optical processor for computing geometrical moments, and demonstrated its use in handwriting recognition, however, the accuracy of this system was not sufficient to accurately compute Hu's moment invariants for most other pattern recognition applications¹⁵.

Many of the hybrid systems described above require coherent input, or the use of a spatial light modulator. This requirement may be too restrictive in some applications. While digital systems provide the flexibility needed in practical applications, the large amount of "number crunching" involved in digital calculations of moment invariants makes real-time digital computation difficult.

An alternative approach is to reduce the amount of information, contained in an input image, that must be processed to determine the value of a moment invariant within the required accuracy. Morris²⁵ has shown that a photon-limited input image can be

used to generate an estimate of the cross correlation between an input image and a reference (or filter) function stored in computer memory. The probabilistic relationship between the spatial coordinates of the detected photoevents and the corresponding location in the input scene provides a natural means of sampling the input. In the next section, we discuss this photon-counting approach, and demonstrate its use in the computation of accurate estimates for radial moments of CHF's.

5.3 Estimation of Radial Moments of CHF's using Photon-Counting Techniques.

In this section we present the theoretical basis for an optical implementation of a Monte Carlo algorithm for the computation of moment invariants which uses a position-sensitive photon-counting detection system. The results depend on a number of probabilistic calculations; these calculations are straightforward, but are somewhat lengthy. Hence, only the results are presented here. The details of the calculations are provided in Appendices A and B. A slightly different definition of the photon-limited signal is required for the estimation of moment invariants. Hence, some of the photon statistics presented earlier are repeated, to provide a clear description of the statistical behavior of this modified correlation signal.

5.3.1 Photon-Counting Methods for Image Recognition

In polar coordinates, a photon-limited input image $f^\dagger(r, \theta)$ can be represented as a two-dimensional collection of Dirac delta functions, i.e.,

$$f^\dagger(r, \theta) = \sum_{i=1}^N \delta(r - r_i, \theta - \theta_i) \quad , \quad (5.10)$$

in which (r_i, θ_i) denotes the spatial coordinates of the i^{th} detected photoevent and N is the total number of detected photoevents. In Eq. (5.10), the spatial coordinates (r_i, θ_i) of the i^{th} detected photoevent are random variables. The number of detected photoevents N may or may not be a random variable, depending on how the experiment is performed. In this treatment, we will take N to be fixed, not random. For this application, the cross-correlation between a photon-limited input scene $f^*(r, \theta)$ and a reference function $R(r, \theta)$ is defined in the following manner:

$$C(r, \theta) = \frac{1}{N} \int \int f^*(r', \theta') R(r + r', \theta + \theta') r' dr' d\theta' \quad (5.11)$$

Using Eqs. (5.10) and (5.11), one obtains

$$C(r, \theta) = \frac{1}{N} \sum_{i=1}^N R(r + r_i, \theta + \theta_i) \quad (5.12)$$

Note that, in previous chapters, the photon-limited correlation signal was realized by sampling the reference function $R(r, \theta)$ at the locations indicated by the detected photoevent coordinates [see Eq. (2.9)]. In this case, the correlation is obtained in a similar fashion, but is **divided** by the number of detected photoevents N . The reason for this definition of the correlation signal will become evident in Section 5.3.3. Once again, the photon-limited correlation signal is a random function, since the event coordinates (r_i, θ_i) are independent random variables with probability density function²⁶

$$P(r_i, \theta_i) = \frac{f(r_i, \theta_i)}{\int_0^{2\pi} \int_0^\infty f(r, \theta) r dr d\theta} \quad (5.13)$$

The mean value of the photon-limited correlation signal is²⁵

$$\langle C(r, \theta) \rangle = \frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') R(r + r', \theta + \theta') r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'} \quad (5.14)$$

where $f(r', \theta')$ is the classical intensity associated with the input image. In Eq. (5.14) note that the mean value of the correlation signal is again directly proportional to the cross correlation between the classical-intensity input image $f(r, \theta)$ and the reference function stored in computer memory. In this case, note that the mean value of the correlation signal given in Eq. (5.14) is independent of the number of detected photoevents N .

To calculate $C(r, \theta)$ in practice, one uses the spatial coordinates of a detected photoevent as an address. The offset coordinate (r, θ) defines the location of the reference function window within the input scene. The procedure is to look up the value of the reference function stored at the address indicated by $(r + r_i, \theta + \theta_i)$, and place the result into an accumulator; this process is repeated for all N detected photoevents. The resulting value in the accumulator is the estimate for the cross correlation between the input scene and the reference function in computer memory.

Since the coordinates of the detected photoevents are independent random variables with the probability density given in Eq. (5.13), the photon-limited correlation signal can be viewed as a Monte Carlo estimate for the cross-correlation integral between the input $f(r, \theta)$ and the reference $R(r, \theta)$. The position-sensitive photon-counting detection system serves as a real-time source of random variates for the Monte Carlo algorithm²⁰.

5.3.2 Estimation of the Input Centroid

Radial moments of CHF's are position-invariant when they are computed about the centroid of the input image. Hence, the first step in computing an estimate for the

moment invariants is to determine the centroid of the input; this is done in the following manner. The Cartesian coordinates of the centroid are computed by summing the x and y coordinates of each detected photoevent into an accumulator, and dividing by the number of detected photoevents N. An alternative way to compute this centroid calculation is to apply, in parallel, the reference functions

$$R_x(x, y) = x \quad (5.15)$$

$$R_y(x, y) = y \quad (5.16)$$

Using Eq. (5.14) (in Cartesian coordinates) the photon-limited correlation signals realized using these reference functions are seen to be an estimate for the x and y coordinates of centroid of the input image. The mean values of the correlation signals obtained using the reference functions in Eqs. (5.15) and (5.16) are given by

$$\langle C_x \rangle = \frac{\iint_A xf(x, y) dx dy}{\iint_A f(x, y) dx dy} \quad , \quad (5.17)$$

and

$$\langle C_y \rangle = \frac{\iint_A yf(x, y) dx dy}{\iint_A f(x, y) dx dy} \quad , \quad (5.18)$$

where A is the area of the input image. Note that Eqs. (5.15) and (5.16) are the exact definitions of the energy centroid of the input scene $f(x, y)$.

By using a sufficient number of detected photoevents, one can compute an estimate for the coordinates of the centroid of the input to within a specified error with a single realization of the correlation signal. The error in the estimate of the centroid that

can be tolerated will depend upon the application in question. The number of detected photoevents required to achieve this error is image-dependent, and can be determined using Eqs. (A31)-(A32) in Appendix A. Once the Cartesian coordinates of the centroid have been determined, that location is chosen as the origin of the polar coordinate system, and all (r, θ) values for subsequently detected photoevents are computed with respect to the centroid.

5.3.3 Estimation of Moment Invariants

Photon-counting methods can be used to provide a Monte Carlo estimate for the moment integral in Eq. (5.3). In this case one takes the reference function stored in computer memory to be

$$R(r, \theta) = r^{s-2} \exp(-im\theta) \quad (5.19)$$

where the coordinates (r, θ) are computed with respect to the centroid of the input image. Using Eqs. (5.14) and (5.19), one finds that the mean value of the subsequent photon-limited correlation signal is

$$\langle C_{s,m} \rangle = \frac{\int_0^{2\pi} \int_0^\infty f(r, \theta) r^{s-1} \exp(-im\theta) dr d\theta}{\int_0^{2\pi} \int_0^\infty f(r, \theta) r dr d\theta}, \quad (5.20)$$

where the offset within the reference window is taken to be the image centroid. (The image centroid can be used as the origin if the input object is the only object present in the input scene, or if the input object has been segmented from the background.) Notice that $\langle C_{s,m} \rangle$ is directly proportional to the classical intensity moment $M_{s,m}$ defined in Eq. (5.3). In addition, if different photons are used to compute $\langle C_{s,m} \rangle$ and

$\langle C_{s,0} \rangle$, the two quantities are statistically independent, and an estimate for $\Phi_{s,m}$ can be obtained from the ratio

$$\frac{|\langle C_{s,m} \rangle|}{|\langle C_{s,0} \rangle|} = \frac{|M_{s,m}|}{|M_{s,0}|} = \Phi_{s,m} \quad (5.21)$$

Equation (5.21) demonstrates how the moment invariants $\Phi_{s,m}$ can be expressed in terms of the mean values of the photon-limited correlation signals. Because of the way in which the photon-limited correlation signal was defined here, the number of detected photoevents is not present in Eq. (5.20), which is advantageous. Clearly, it is desirable to obtain an estimate for $\Phi_{s,m}$ from a single realization of the appropriate correlation signals, rather than perform enough realizations to determine the expected value. Unfortunately, for arbitrary m and s this is not possible. As detailed in Appendix B, only the large, or dominant moments of a given image can be accurately estimated using a single realization of the photon-limited correlation signal. However, the strong moments are typically the useful ones for pattern recognition applications. Using the theory presented in Appendices A and B, one can predict in advance whether a given moment can be accurately estimated using photon-counting techniques. As indicated in Section 5.2.2, the low-order moments are the most important ones for pattern recognition. As demonstrated later in Section 5.4.1, in Tables 5.1-5.3, the strong moments in the set $\Phi_{s,m}$ ($s=1$ to 3, $m=1$ to 9) are estimated accurately using photon-counting techniques.

5.3.3.1 Estimation of $\Phi_{2,m}$

The moment invariant $\Phi_{2,m}$ is particularly suited to estimation using photon-counting techniques, and is the radial-order moment that is most often used in the literature for pattern recognition purposes. If we let $s=2$ in Eq. (5.19), then using Eq. (5.14), the modulus of the mean value of the photon-limited correlation is

$$|< C_{2,m} >| = \left| \frac{\int_0^{2\pi} \int_0^{\infty} f(r, \theta) \exp(-im\theta) r dr d\theta}{\int_0^{2\pi} \int_0^{\infty} f(r, \theta) r dr d\theta} \right| = \Phi_{2,m} \quad (5.22)$$

In Eq. (5.22), an important feature to note is that the normalization by the 0th - order angular moment $|M_{2,0}|$ occurs naturally, as a result of the statistics of the detected photoevents. In this case, it was not necessary to compute $|M_{2,m}|$ and $|M_{2,0}|$ separately by determining $<|C_{2,m}|>$ and $<|C_{2,0}|>$. Hence, with a single reference function, one can determine an estimate value for $\Phi_{2,m}$, whereas with the digital computation of $\Phi_{2,m}$ both $|M_{2,m}|$ and $|M_{2,0}|$ must be computed individually.

It is desirable to obtain an accurate estimate for $\Phi_{2,m}$ using a single realization of the modulus of the correlation signal $|C_{2,m}|$. A necessary condition is that the mean value of the estimate be approximately equal to the moment in question; i.e.

$$<|C_{2,m}|> \cong \Phi_{2,m} \quad (5.23)$$

Fortunately, as shown in Appendix B, and in Section 5.4 (Tables 5.1-5.3), Eq. (5.23) is approximately correct for the strong moments of real input images. In addition, it is possible in many cases to obtain an estimate for $\Phi_{2,m}$ even when the moment that is being estimated is not strong. This can be obtained by using

$$\left(|C_{2,m}|^2 - \frac{1}{N} \right)^{\frac{1}{2}} \quad (5.23a)$$

as the estimate for $\Phi_{2,m}$ (the rationale for this is given in Appendix B).

Hence, the procedure for determining an estimate for $\Phi_{2,m}$ is as follows:

- 1.) Compute photon-limited correlation signals C_x and C_y using the reference functions in Eqs. (5.15) and (5.16) to

determine estimates for the Cartesian coordinates of the centroid of the input to within the necessary error. (The number of photons required to achieve this accuracy can be predicted using Eqs. (A25)-(A26) for a typical input image.)

2.) Use the centroid coordinates as the origin for the polar coordinate system, and compute (r_i, θ_i) values for the next N detected photoevents. (The procedure for determining N for a given application is discussed in Section 5.3.4.)

3.) The photon-limited estimate for $\Phi_{2,m}$ is obtained by computing the modulus of $C_{2,m}$ using Eqs. (5.12) and (5.19), where $s=2$, and the offset is taken to be zero. The estimate is then determined using relation (5.23a)

5.3.3.2 Estimation of $\Phi_{1,m}$

To estimate $\Phi_{1,m}$, one must compute estimates for $|M_{1,m}|$ and $|M_{1,0}|$. If one chooses the reference functions

$$R_{1,m} = r^{-1} \exp \{ -im\theta \} \quad (5.24)$$

and

$$R_{1,0} = r^{-1} \quad , \quad (5.25)$$

one can obtain an estimate for $\Phi_{1,m}$ by computing $C_{1,m}$ and $C_{1,0}$ using Eq. (5.12). The estimate for $\Phi_{1,m}$ is obtained by taking the ratio of the moduli of the two correlation signals.

5.3.4 Determining the Required Number of Detected Photoevents N

The number of detected photoevents N required to produce an accurate estimate for a desired moment invariant can be specified in one of two ways. The most obvious

way is to require that a given moment invariant be determined to within a specified error. One measure of this error is the ratio of the standard deviation to the mean of the estimate be less than a given value. One might require that

$$\frac{\sigma_{s,m}}{\langle \Phi'_{s,m} \rangle} < E, \quad (5.26)$$

where $\Phi'_{s,m}$ is the photon-limited estimate for the moment invariant $\Phi_{s,m}$, $\sigma_{s,m}$ is the standard deviation of the estimate, and E is the error that can be tolerated in the estimate for $\Phi_{s,m}$.

As an example, we outline the procedure for determining the number of detected photoevents that will produce an estimate for $\Phi_{2,m}$ to within given accuracy for a given input. First, one determines mean values and standard deviations for the real and imaginary parts of the photon-limited correlation signal obtained using the reference function in Eq. (5.19) with $s=2$, for an initial number of detected photoevents N . The mean values, standard deviations and correlation coefficients are obtained using Eqs. (A6)-(A10) in Appendix A. Next, these values are substituted into Eq. (A15) to obtain an expression for the probability density function for the estimate of $\Phi_{2,m}$. The mean values and standard deviations for the estimate of $\Phi_{2,m}$ are obtained (see Appendix A) using the density function for the estimate of $\Phi_{2,m}$, from which the accuracy of the estimate is determined using Eq. (5.26). If the accuracy is not sufficient, the above process is repeated for increasing numbers of photoevents until the desired accuracy is achieved.

An alternative way to specify the number of detected photoevents is to require that only enough photoevents be detected for the estimates for the moments to allow one to distinguish among a given set of images. The accuracy required for the estimate of a given set of moment invariants depends on the images that are to be discriminated

among. The recognition decision criterion is the distance in feature space D^2 [defined in Eq. (5.9)] between the moment invariants for the input and reference images. The more similar the moment invariants are for a given set of images, the smaller the distance in feature space D^2 will be. The probabilistic nature of the photon-counting technique for estimation of moment invariants results in a statistical spread in the estimates for D^2 . Equations (A8) and (A9) of Appendix A show that the variance of the photon-limited correlation signal decreases as $1/N$, where N is the number of detected photoevents. Hence, the variance in the estimates for the moment invariants and the distance D^2 will also decrease with increasing N .

The exact number of detected photoevents required to discriminate among a set of given images can once again be determined using the statistical theory of hypothesis testing²⁷. On the basis of the photon-limited Monte Carlo estimate for D^2 , one must choose between two hypotheses: the null hypothesis H_0 -- input image $f(r,\theta)$ is *not* the same as the reference image $R(r,\theta)$, or the positive hypothesis, H_1 -- the input image $f(r,\theta)$ is the same as the reference image $R(r,\theta)$. Under hypothesis H_0 , the probability density function for D^2 is denoted by $P_0(D^2)=P[D^2 | f(r,\theta)=N(r,\theta)]$, where $N(r,\theta)$ is some noise, or false image. Under Hypothesis H_1 , the density function of D^2 is denoted by $P_1(D^2)=P[D^2 | f(r,\theta)=R(r,\theta)]$, where $R(r,\theta)$ is the reference image.

As indicated in Section. 5.2.2, the observer sets a threshold D_T for the distance in feature space. If the estimate for the distance $D^2 > D_T$, hypothesis H_0 is chosen; if $D^2 < D_T$, hypothesis H_1 is chosen. However, because of the statistical nature of the estimate for D^2 , the observer occasionally makes an error, regardless of the value chosen for D_T . The probability of choosing H_1 when H_0 is true is called the probability of false alarm and is given by

$$P_{fa} = \int_{D_T}^{\infty} P_0(D^2) d(D^2) \quad . \quad (5.27)$$

The probability of choosing H_1 when H_1 is true is called the probability of detection, and is given by

$$P_d = \int_{D_T}^{\infty} d(D^2) P_1(D^2) \quad . \quad (5.28)$$

To determine the number of detected photoevents N that are required to calculate each moment invariant within a given accuracy, one must specify the required probabilities of detection and false alarm for a given application. Next, one calculates the probability density functions for the photon-limited estimates for D^2 for the input and reference images for a starting number of detected photoevents N ; the functional form of this density function is given in Eq. (A26). One then computes P_d and P_{fa} using Eqs (5.27) and (5.28), and compares these values with the required ones. If the calculated values for P_d and P_{fa} do not meet the desired specifications, one increases the value for N , and repeats the process until the required probabilities of detection and false alarm are achieved.

One should note several points regarding the above process. First, all of the computations of the various density functions can be done off-line, before the recognition system is actually employed. Hence, the time required to perform these computations is not critical. Second, it is not necessary to determine the required number of detected photoevents exactly. If a few hundred more photoevents than necessary are used to compute the estimates for D^2 , only a few tens of milliseconds are added to the time required to compute an estimate (assuming the detection system is operating at ~ 100 kHz). Finally, in practice it is probably easier to determine an

approximate value for the necessary number of photoevents through a brief trial-and-error process. That is, once the moments for the reference image have been computed, the mean value and standard deviation for the distance in feature space can be easily approximated experimentally by performing one hundred realizations of the photon-limited estimates for the distance in feature space for each input image (100 realizations takes about five seconds). This can be done for two or three different photon levels, which is usually sufficient for determining the number of photons necessary to achieve the error given in Eq. (5.26). However, if one has specified the acceptable error in terms of P_d and P_{fa} , one must still calculate the density function given in Eqs. (A25) and (A26) for each input image.

5.4 Experimental Results

Two different sets of experiments were performed to test the theoretical predictions given in Section 5.3. In the first set, experiments were performed that compared the theoretical predictions with the experimental values for $\Phi_{1,m}$, $\Phi_{2,m}$ and $\Phi_{3,m}$, for $m=1$ to 9. This was done using the classical-intensity input objects shown in Fig. 5.1. The second set, in which the goal was to discriminate among the objects in Fig. 5.1, tested the recognition capabilities of the moment invariants.

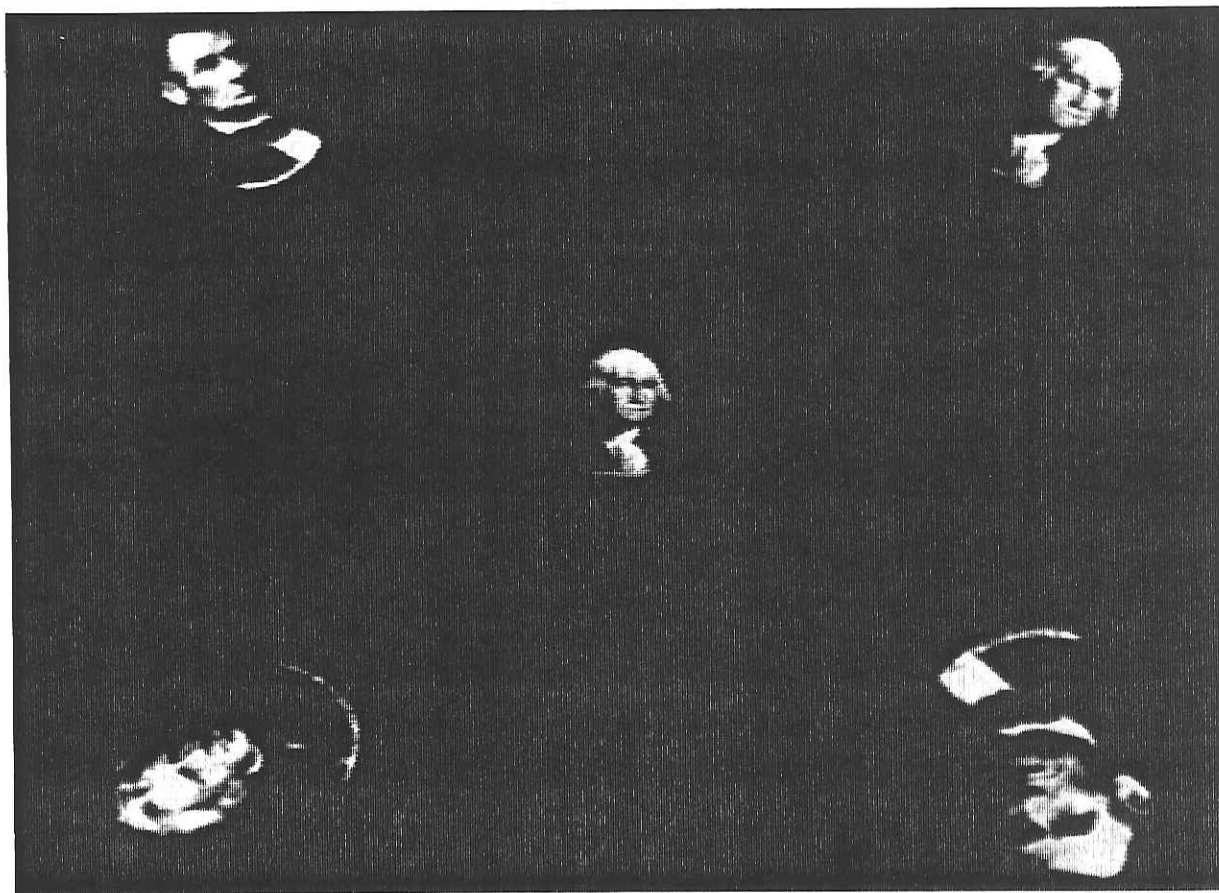


Figure 5.1 Engraved portrait objects used in image recognition experiments.

5.4.1 Estimation of Moment Invariants

The experimental system is shown schematically in Fig. 5.3. The input objects (in this case images of presidents on U. S. currency) were illuminated with incoherent light, and imaged onto a two-dimensional, photon-counting detector²⁸. Position-computing electronics²⁹ calculate (in real time) the (x,y) coordinates of the detected photoevents with eight-bit resolution; these coordinates are then sent to a digital computer for processing. Sufficient neutral density is inserted between the input and the detector to achieve the necessary low-light-level conditions at the detector. In these experiments, the illumination was provided by fluorescent room lights, and a neutral density of four was used to achieve a detection rate of 50 kHz. (The maximum count rate for the detector is 100 kHz.)

The procedure outlined in Section 5.2 was used to compute estimates for $\Phi_{1,m}$, $\Phi_{2,m}$ and $\Phi_{3,m}$ for the classical-intensity input objects shown in Fig. 5.1, where the values of m ranged from one to nine. For each input object, 5000 photoevents were detected to determine the x and y coordinates of the centroid using the reference functions given by Eqs. (5.15) and (5.16). This resulted in a standard deviation in the estimate for the centroid coordinates of 0.25 pixels in each spatial dimension. Next, 5000 more detected photoevents were used to evaluate the photon-limited correlation signals $C_{s,m}$ using Eq. (5.12), using the appropriate values for m and s in the reference function given in Eq. (5.19). The estimates for $\Phi_{1,m}$ and $\Phi_{3,m}$ were obtained by forming the ratios $|C_{1,m}|/|C_{1,0}|$ and $|C_{3,m}|/|C_{3,0}|$ respectively. The estimates for $\Phi_{2,m}$ were obtained using $|C_{2,m}|$ [see Eqs. (5.22) and (5.23)].

One thousand realizations for each moment were computed to verify the predicted statistical behavior of the estimates for the moment invariants. The objects

were input at various orientations ranging between 0 and 360 degrees, with relative scales ranging from 1.0-2.0 (see Fig. 5.1).

Representative results for the case when Lincoln was input are shown in tabular form in Tables 5.1-5.3. The high-light-level moments are computed using Eqs. (5.3) and (5.4). The theoretical mean values for the estimates of $\Phi_{2,2}$ and $\Phi_{2,4}$ are computed numerically using the probability density function given in Eq. (A15). The theoretical values for the estimates of $\Phi_{2,m}$, $m \neq 2$ or 4, are computed using Eq. (B15). The theoretical values of $\Phi_{1,m}$ and $\Phi_{3,m}$ are computed using the probability density function given in Eq. (A19).

The estimate for the moment invariant is said to be accurate if the following two conditions are satisfied. First, the mean value of the estimate must be approximately equal to the value of the moment computed from the high-light-level image using Eqs. (5.3) and (5.4). Second, the ratio of the standard deviation of the estimate to the expected value of the estimate must be small [see Eq. (5.26)]. Examination of Tables 5.1 and 5.3 clearly shows that for $s=1$ or 3, only the stronger moments can be estimated accurately using photon-counting techniques. In Table 5.1, only the moment $\Phi_{1,2}$ is estimated with any reliability using 5000 detected photoevents. Because the mean value of the photon-limited estimate for the other moments is not approximately equal to the value of the high-light-level moments, the other order moments cannot be accurately estimated using photon-counting techniques. In Table 5.3, the moments $\Phi_{s,m}$, $s=3$; $m=2, 3, 4$ and 6 are the most accurately estimated. Note that the two strongest moments, $\Phi_{3,2}$ and $\Phi_{3,4}$, are estimated with the least amount of error and the greatest reliability. Table 5.2 demonstrates that $\Phi_{2,m}$ are estimated with the least error and best reliability.

Fortunately, one can predict in advance which moments can be estimated using photon-counting techniques by computing numerically the mean values of the photon-limited estimates. The reason for the difficulty in estimating weak moments is described in Appendix B. However, the useful moments for pattern recognition purposes are typically the strong moments; hence this inability to compute the weak moments accurately is not a significant limitation in pattern recognition applications.

One thousand realizations of each moment were performed to test the theoretical predictions for the PDF's of the estimates. As an example, histograms of the estimates for the moments $\Phi_{2,2}$ and $\Phi_{2,4}$ of Lincoln, Washington and Jackson are shown in Figs. 5.3, 5.4 and 5.5, respectively. The solid lines indicate the theoretical predictions for the density functions made using Eq. (A15). In each case, the engraved portraits were input at unit magnification.

Order	High-light- level moments	Theoretical Mean	Experimental Mean	Sigma	% Error	Sigma/Mean
s=1; m						
1	.598E-01	.656E-01	.767E-01	.161E-01	16.9	0.21
2	.752E-01	.733E-01	.781E-01	.149E-01	3.9	0.19
3	.123E-01	.192E-01	.432E-01	.142E-01	251	0.33
4	.567E-01	.583E-01	.622E-01	.179E-01	9.7	0.29
5	.151E-01	.221E-01	.272E-01	.117E-01	80.1	0.43
6	.238E-01	.252E-01	.287E-01	.152E-01	20.6	0.53
7	.127E-01	.165E-01	.169E-01	.863E-02	33.1	0.51
8	.136E-01	.168E-01	.299E-01	.116E-01	120	0.39
9	.709E-02	.236E-01	.224E-01	.954E-02	215	0.43

Table 5.1 Comparison of high-light-level moments and photon-limited estimates for $\Phi_{1,m}$ m=1 to 9.

The input image is Lincoln. The theoretical mean is the theoretical prediction for the expected value of the photon-limited estimate. The experimental mean is the experimentally observed expected value for the photon-limited estimate.

The % error is defined as (High-light-level moment-Experimental Mean)/High-light-level moment x 100%.

Order s=2; m	High-light- level moments	Theoretical Mean	Experimental Mean	Sigma	% Error	Sigma/Mean
1	.211E-01	.156E-01	.155E-01	.981E-02	-27.1	0.63
2	.117E+00	.118E+00	.119E+00	.863E-02	1.2	0.07
3	.645E-01	.655E-01	.660E-01	.109E-01	2.3	0.17
4	.131E+00	.128E+00	.127E+00	.981E-02	-3.0	0.08
5	.218E-01	.293E-01	.269E-01	.998E-02	23.0	0.37
6	.523E-01	.535E-01	.526E-01	.888E-02	0.6	0.17
7	.292E-01	.324E-01	.297E-01	.101E-01	1.7	0.34
8	.319E-01	.360E-01	.344E-01	.101E-01	7.8	0.30
9	.277E-01	.278E-01	.241E-01	.975E-02	13.0	0.40

Table 5.2 Comparison of high-light-level moments and photon-limited estimates for $\Phi_{2,m}$ m=1 to 9.

The input image is Lincoln. The theoretical mean is the theoretical prediction for the expected value of the photon-limited estimate. The experimental mean is the experimentally observed expected value for the photon-limited estimate.

The percent error is defined as: (High-light-level moment-Experimental Mean)/High-light-level moment x 100%.

Order	High-light- level moments	Theoretical Mean	Experimental Mean	Sigma	% Error	Sigma/Mean
s=3; m						
1	.367E-01	.141E-01	.123E-01	.978E-02	-66.2	0.8
2	.150E+00	.149E+00	.151E+00	.981E-02	0.7	0.06
3	.805E-01	.850E-01	.867E-01	.991E-02	7.7	0.11
4	.206E+00	.195E+00	.191E+00	.979E-02	-7.3	0.05
5	.333E-01	.398E-01	.437E-01	.984E-02	31.2	0.22
6	.858E-01	.881E-01	.837E-01	.993E-02	-2.4	0.12
7	.441E-01	.480E-01	.493E-01	.987E-02	11.8	0.19
8	.497E-01	.447E-01	.527E-01	.989E-02	6.0	0.30
9	.343E-01	.336E-01	.322E-01	.969E-02	-6.1	0.40

Table 5.3 Comparison of high-light-level moments and photon-limited estimates for $\Phi_{3,m}$ $m=1$ to 9.

The input image is Lincoln. The theoretical mean is the theoretical prediction for the expected value of the photon-limited estimate. The experimental mean is the experimentally observed expected value for the photon-limited estimate.

The % error is defined as (High-light-level moment-Experimental Mean)/High-light-level moment x 100%.

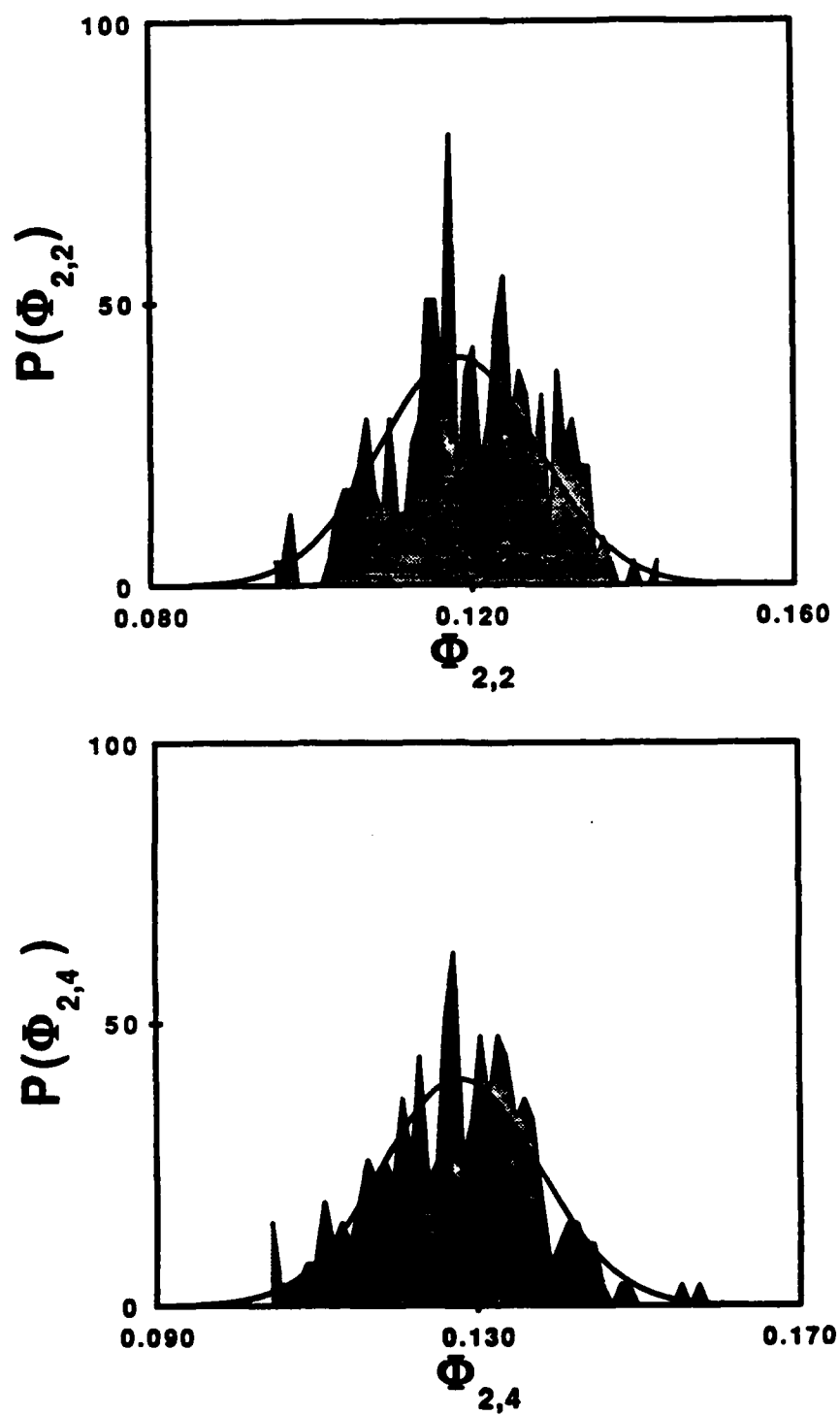


Fig. 5.2 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Lincoln. The solid lines indicate theoretical predictions for the probability density function.

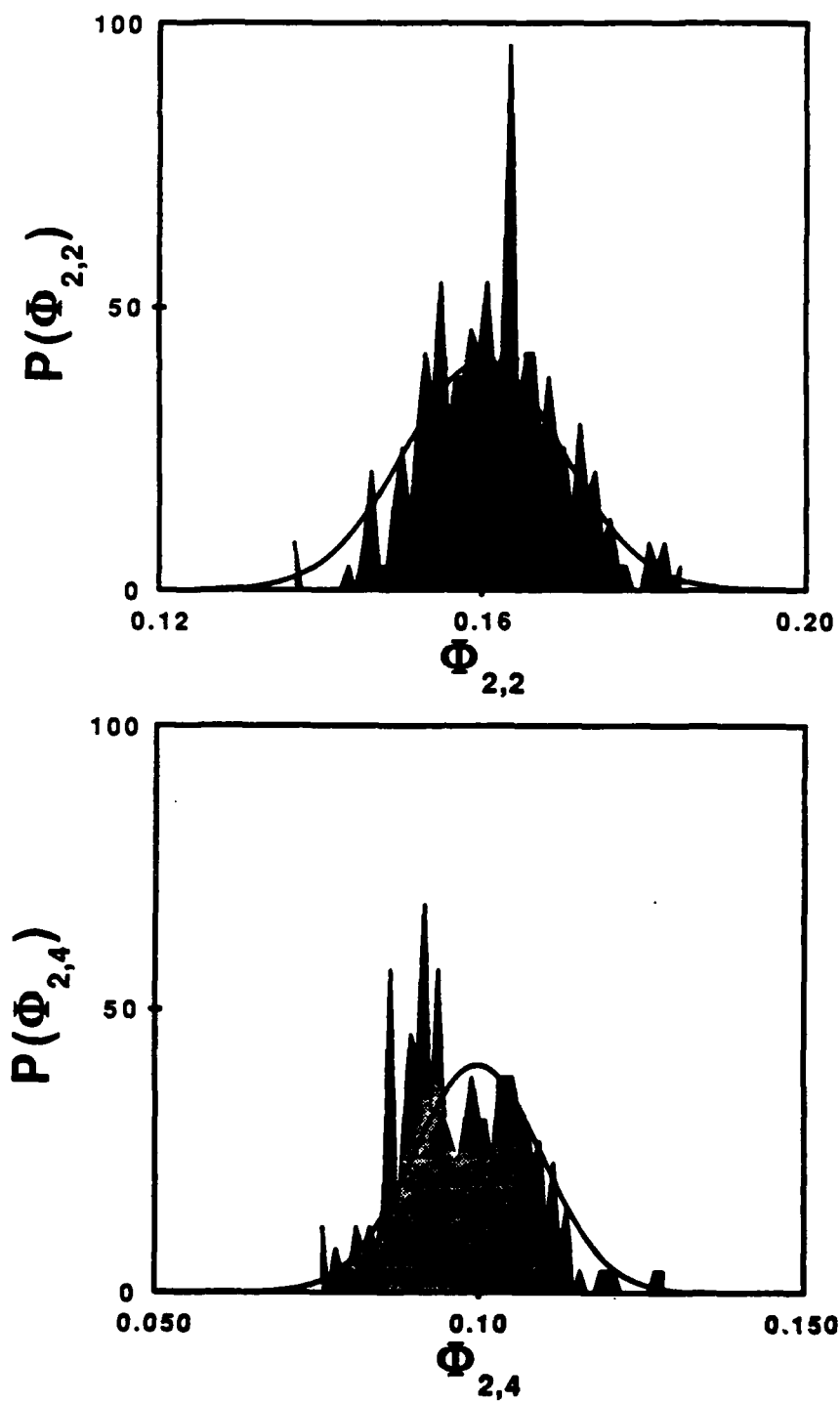


Fig. 5.3 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Washington. The solid lines indicate theoretical predictions for the probability density function.

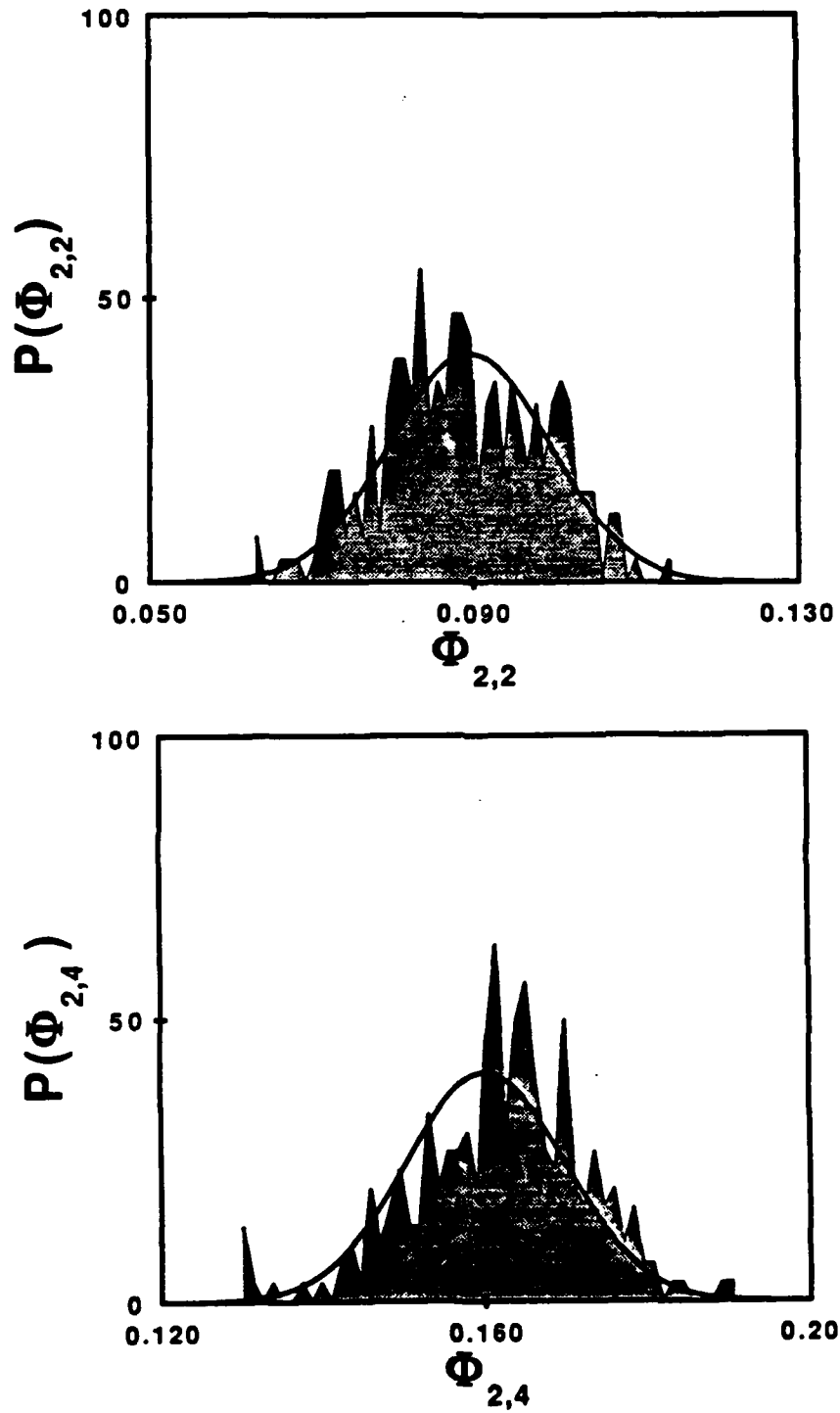


Fig. 5.4 Histogram of experimental values obtained for $\Phi_{2,2}$ and $\Phi_{2,4}$ when the input image is Jackson. The solid lines indicate theoretical predictions for the probability density function.

5.4.2 Pattern Recognition Performance

In this set of experiments, we test the pattern recognition performance of the photon-limited estimates of the moment invariants. After examining the relative strengths of the various moments for the the high-light-level objects shown in Fig. 5.2, it was determined that the use of moment invariants $\Phi_{2,2}$ and $\Phi_{2,4}$ would provide a separation in feature space [see Eq. (5.9)] large enough to allow the portrait of Lincoln to be distinguished from the portraits of Washington and Jackson. Using the method of hypothesis testing outlined in Section 5.3.4, it was determined that 5000 detected photoevents would provide accurate estimates for the moment invariants such that the probability of making an incorrect decision was less than 1×10^{-3} . This probability of error was achieved using a distance threshold of .02 in feature space. This threshold was determined by plotting the probability density functions for the distance in feature space between the reference object and the input objects using Eqs. (A29) and (A20). The PDF's were plotted using an increasing number of detected photoevents until the above probability of error was obtained.

Photon-limited estimates for the moment invariants $\Phi_{2,2}$ and $\Phi_{2,4}$ were computed for each input object. The distances in feature space between the reference moments for Lincoln that are stored in computer memory and the moments for the input images were then computed, and compared with the decision threshold. If the distance was less than the threshold, the input object was identified as Lincoln. To test the theoretical predictions for the probability density function for the distance in feature space [see Eq. (A19)], 1000 determinations of the distance in feature space were performed for each input object. As in Section 5.4.1, the input objects were input at various orientations, and at relative scales ranging from 1.0 to 2.0. The results are shown in histogram form in Figs. 5.5 and 5.6 in which the solid lines indicate the

theoretical predictions for the PDF's of the photon-limited estimates of the distance in feature space obtained using Eqs. (A19) and (A20). The theoretical predictions were based on images input at a relative scale of 1.5. There is a 7 percent change in the theoretical values of the moment invariants for images with a relative scale of 1.0, versus images with a scale of 2.0. This is due to an inherent quantization error and is one of the limiting characteristics of this method for image recognition. When the range of scales is known, the reference moments are usually chosen for a reference object whose size is in the middle of the range, as was done here. This accounts for the majority of the error between theory and experiment shown in Figs. 5.5 and 5.6. Note that even in the presence of this quantization error, it is still possible to set a discrimination threshold such that the reference images can still be recognized with a probability of making an incorrect decision on the order of 1×10^{-3} .

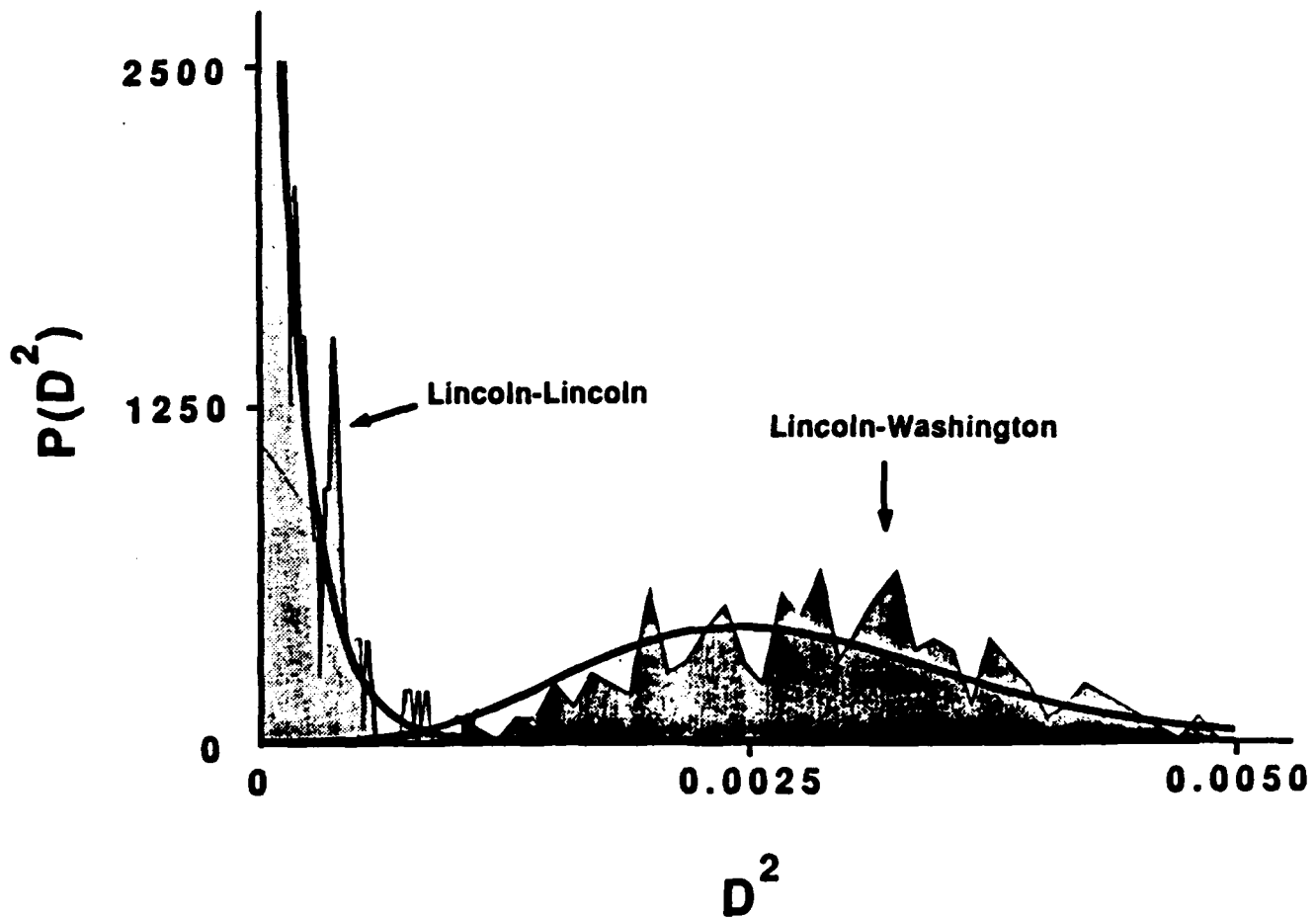


Figure 5.5 Histograms of experimentally obtained estimates for the distance in feature space D^2 between the portraits of Lincoln and the portraits of Washington, using the features $\Phi_{2,2}$ and $\Phi_{2,4}$. The solid lines indicate theoretical predictions for the probability density functions of the distance in feature space.

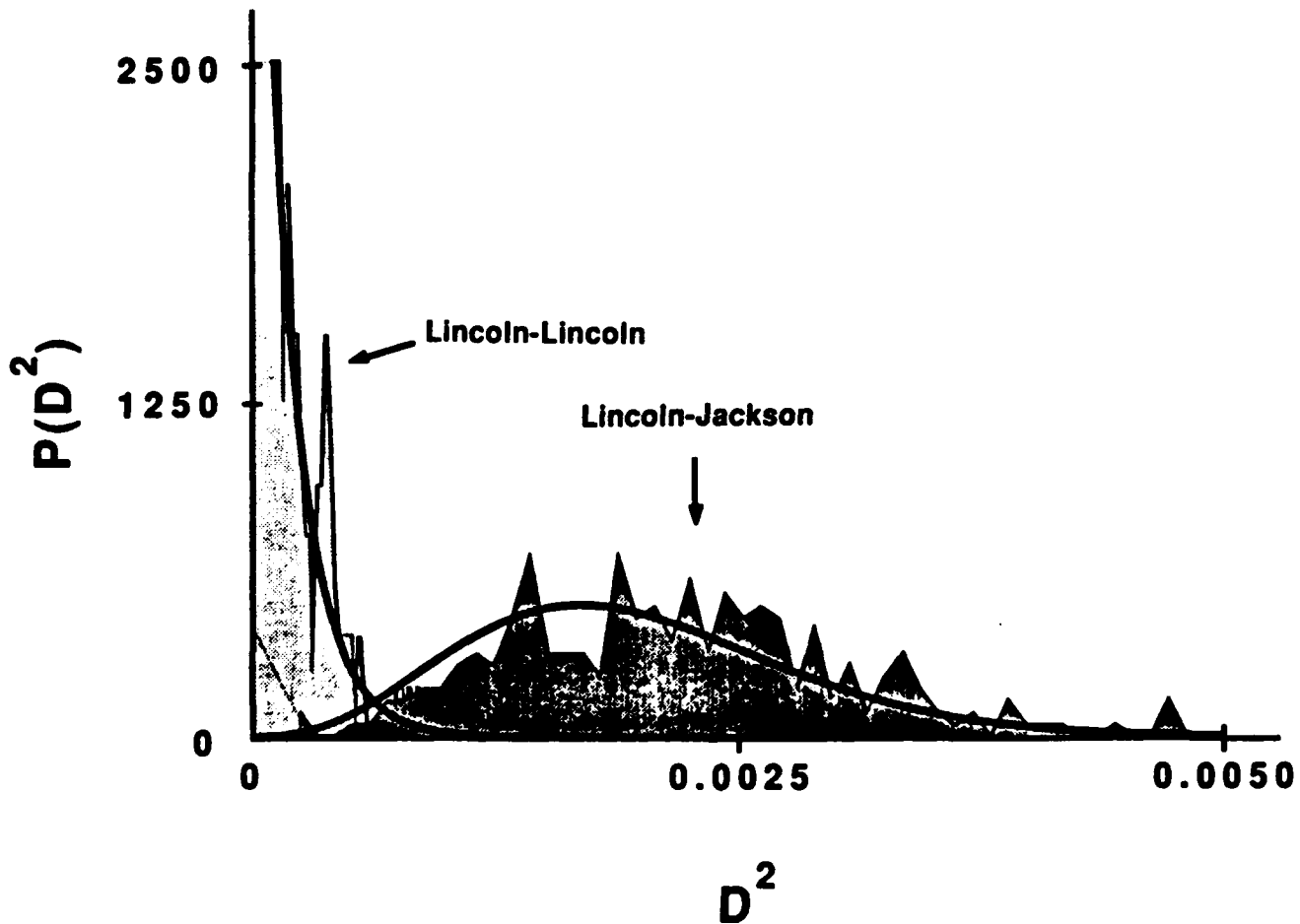


Figure 5.6 Histograms of experimentally obtained estimates for the distance in feature space D^2 between the portraits of Lincoln and the portraits of Jackson, using the features $\Phi_{2,2}$ and $\Phi_{2,4}$. The solid lines indicate theoretical predictions for the probability density functions of the distance in feature space.

5.5 Summary

In this chapter, the estimation of radial moments of circular-harmonic functions using photon-counting techniques is investigated. In Section 5.2, the basic method of moment invariants for pattern recognition is briefly reviewed. In Section 5.3, a method for the real-time estimation of the moment invariants using the position-sensitive photon-counting detection system is presented. The statistical description of the estimates is presented in Appendix A. In Section 5.4, experimental confirmation of the theory is presented. A comparison of theory and experiment for the moment estimates is given in Tables 5.1-5.3, and in Figs. 5.2-5.4. In addition, a pattern recognition application is considered in which engraved portrait images are input at all orientations, and at relative scales that vary up to a factor of two. The results are given in Figs. 5.5 and 5.6. Finally, some issues regarding the computation of weak image moments are discussed in Appendix B.

References for Chapter 5

1. M.K. Hu, 'Visual pattern recognition by moment invariants,' IRE Trans. Inf. Theory IT-8, 179-187 (1961).
2. R.Y. Wong and E.L. Hall, 'Scene matching with invariant moments,' Comput. Graphics Image Process. 8, 16-21 (1978).
3. S.A. Dudani, K.J. Breeding and R.B. McGee, 'Aircraft identification by moment invariants,' IEEE Trans. Comput. C-26, 39-45 (1977).
4. S.Maitra, 'Moment invariants,' Proc. IEEE 67, 697-699 (1979).
5. A. Gostasby, 'Template matching in rotated images,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-7, 338-343 (1985).
6. D. Casasent and D. Pauly, 'Infrared ship classification using a new pattern recognition concept,' Proc. Soc. Photo-Opt. Instrum. Eng. 302, 126-133 (1981).
7. Y. Abu-Mostafa and D. Psaltis, 'Recognitive aspects of moment invariants,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-6, 698-706 (1984).
8. Y. Abu-Mostafa and D. Psaltis, 'Image normalization by complex moments,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-7, 46-55 (1985).
9. C. H. Teh and R. T. Chin, 'On digital calculations of moment invariants,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-7, 640-641 (1985).
10. D.Casasent and D. Psaltis, 'Hybrid processor to compute invariant moments for pattern recognition,' Opt. Lett. 5, 395-398 (1980).
11. B.V.K. Kumar and C.A. Rahencamp, 'Calculation of geometric moments from Fourier plane intensities,' Appl. Opt. 6, 997-1007 (1986).
12. K. Wagner and D. Psaltis, 'Real-time computation of moments with acousto-optics,' Proc. SPIE 352, 82-89 (1982).
13. M. R. Teague, 'Optical calculation of irradiance moments,' Appl. Opt. 19, 1353-1358 (1980).
14. R.A. Athale and J. A. Blodgett, 'Coherent optical moment generator with improved performance,' J. Opt. Soc. Am. 72, 1779-1783 (1982).
15. J. Duvernoy and Y. Sheng, 'Effective optical processor for computing image moments at TV rate: use in handwriting recognition,' Appl. Opt. 26, 2320-2327 (1987).
16. Y.N. Hsu and H.H. Arsenault, 'Rotation-invariant pattern recognition using circular harmonic expansion,' Appl. Opt. 22, 130-132 (1983).

17. Y.N. Hsu and H.H. Arsenault, 'Experiments on pattern recognition using invariant Fourier-Mellin descriptors,' J. Opt. Soc. Am. A 3, 771-776 (1986).
18. Y. Sheng and J. Duvernoy, 'Circular-Fourier-Radial-Mellin transform descriptors for pattern recognition,' J. Opt. Soc. Am. A 3 885-888 (1986).
19. M. Lampton and C. W. Carlson, 'Low-distortion resistive anodes for two-dimensional position-sensitive MCP systems,' Rev. Sci. Instrum. Eng. 331, 365-371 (1982).
20. M. H. Kalosh and P. A. Whitlock, *Monte Carlo Methods, Vol. 1* (Wiley, New York, 1986).
21. G.M. Morris, 'Optical computing using Monte Carlo methods,' Opt. Eng. 24, 86-90 (1985).
22. T.A. Isberg and G.M. Morris, 'Rotation-invariant image recognition at low light levels,' J. Opt. Soc. Am. A 3, 954-963 (1986).
23. M.N. Wernick and G.M. Morris, 'Image classification at low light levels,' J. Opt. Soc. A 3, 2179-2187 (1986).
24. S. S. Reddi, 'Radial and angular moment invariants for image identification,' IEEE Trans. Pattern Anal. Machine Intel. PAMI-3, 240-247 (1981).
25. G. M. Morris, 'Scene matching at low light levels: a computer simulation,' Appl. Opt. 23, 3152-3159 (1984).
26. M. Bertolotti, 'Photon Statistics,' in Photon Correlation and Light Beating Spectroscopy, H. Z. Cummins and E. L. Pike, eds. (Plenum, New York, 1974), Chap. 2.
27. C.W. Helstrom, *Statistical Theory of Signal Detection* (Pergamon, Oxford, 1968), Chap. 3.
28. ITT Electro-Optical Products Division, Tube and Sensors Laboratories, Fort Wayne, IN. Detector model F4146M.
29. Surface Science Laboratories Inc., Mountain View, CA. Model 2401 position computer.

Chapter 6

Two-Stage Template Matching using Quantum-Limited Images

6.1 Introduction

Template matching is a classical approach to the problem of locating and identifying a particular object from within an input scene¹⁻¹². In this technique for image recognition, a large input scene is searched for a particular reference object(s) by applying a small template (or reference function) at each offset location in the input scene. Some measure of the similarity between the template and the input scene at each location is calculated. Based on the similarity measure, it is determined whether the reference object is present at a given location. Obviously, computing the similarity criterion at every possible offset in the input scene can be a very time consuming process, even when a small template is used. This has been a major limitation to this method for image recognition.

To reduce the amount of computation involved, several algorithms have been developed^{2-5,7}. One particular algorithm is a two-stage method of template matching^{4,5,7}. In the first stage, only a small number of points in the input scene that fall within the reference window are randomly sampled and used to compute an estimate for the similarity criterion at each reference window offset. In the second stage, locations that satisfy the similarity criterion in the first stage are examined using all of the points in the input scene that lie within the reference window at those

locations. In essence, the first stage amounts to a coarse search of the input scene to determine likely locations for the reference object. This is equivalent to the 'first look' frequently referred to in the target recognition literature¹³. The number of locations searched in the second stage will depend on the threshold set for the similarity criterion in the first stage. The goal is to minimize the amount of computation involved in the implementation of the two-stage process.

In this chapter, we demonstrate how this two-stage template matching process can be implemented in a highly-efficient manner using the photon-counting detection system described in earlier chapters. A high-light-level input scene is searched using a small window function that is moved to various offsets within the input scene. The cross-correlation between the reference function and the input scene at a given offset is estimated by sampling the input scene using a small number of detected photoevents. The photon-limited correlation signal is used as a similarity criterion for comparing the input with the reference object. In the first stage, a small number of detected photoevents is used to find probable locations for the reference object. In the second stage, likely locations are examined using a sufficient number of detected photoevents to provide a recognition decision within specified probabilities of detection and false alarm.

In Section 6.2, the two-stage template matching process is reviewed. In Section 6.3, the theoretical formalism is presented for the implementation of the two-stage template matching technique at low light levels. In Section 6.4, computer simulations are presented that demonstrate the performance of the two-stage template matching process using a photon-counting detection system. Results are presented for a simple inspection problem, for scene matching in a satellite image, and for automatic

target recognition in a cluttered environment. Finally, in Section 6.5, the rotation-invariant circular-harmonic filter described in Chapter 4 is also employed using this two-stage filtering technique.

6.2 Two-Stage Template Matching

Template matching has been used as a method for locating a particular sub-image from within a larger input scene (see Fig.6.1) for almost 30 years¹. The reason for locating the sub-image may be for the purpose of pattern recognition, where one desires to identify and locate a particular object within an input scene. Alternative applications for template matching are in the registration of digital images, (particularly satellite images), and as a pre-screen method for photo-interpreters. The template matching technique can be employed to identify locations of interest in an aerial photograph or satellite image, and the final recognition decision is made by a human photo-interpreter. In most cases, a large input scene is searched by comparing various locations within the scene to a small reference function. One defines a measure of the similarity between the reference function and input scene at a given offset; when the similarity criterion exceeds some pre-determined threshold, the input and reference are said to be the same. Hence, depending on the application, the object of interest has been detected, or the input scene is said to be registered.

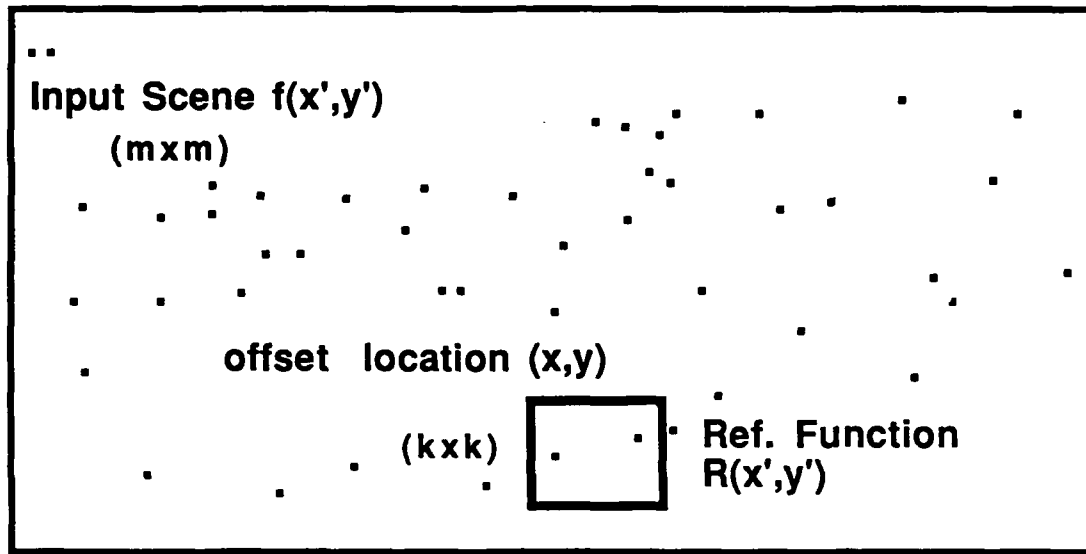


Figure 6.1. Quantum-limited input scene showing reference function window at offset location (x,y) .

Similarity criteria include normalized cross-correlation,⁵ the sum of absolute differences between the pixels⁴, and other methods². Obviously, it is a very computationally intensive process to compute the similarity measure at each location, even if the reference function that is employed is small. Several algorithms have been developed^{2-5,7} that offer the potential to substantially reduce the amount of computation involved in the template matching process. An algorithm that has proven to be quite successful is the two-stage method of template matching^{4,5,7}.

As mentioned earlier, the two-stage template matching algorithm involves moving a reference window to each offset location in the input scene. At each location, a small number of points are sampled, and are used to estimate the similarity criterion

between the input and reference at that offset. If the similarity criterion exceeds the threshold, that location is examined fully in the second stage. In the past⁴, the order in which the points were sampled were chosen at random, using a random number generator. The goal of the first stage is to identify locations that are likely to contain the reference object(s) of interest. The threshold for the similarity criterion must be chosen such that all of the target locations will exceed the threshold, with as few false alarms as possible.

In some cases, it possible to reduce the number of locations that must be examined in the second stage by pre-processing the input scene before beginning the first stage. Such pre-processing techniques include segmenting objects of interest from the background (e.g. using Hough transform techniques¹¹), or reducing the number of gray levels in the input. Whether it is advantageous, or possible to implement a pre-processing step must be considered on a case-by-case basis. In either event, the goal is to minimize the amount of computation involved in the implementation of the two-stage template matching process.

To determine the threshold for the similarity criterion for each stage, it is useful to define a cost function, K , for the entire template matching process, where K is given by

$$K=(m-k)^2 N_1 + N_E N_2 \quad (6.1)$$

In Eq. (6.1), m is the dimension of the input scene, k is the dimension of the template (see Figure 6.1), N_1 is the number of points that are sampled at each location in the first stage, N_E is the number of locations that exceeded the threshold in the first stage, and N_2 is the number of points that are sampled at each location in the second stage. The

parameters available for use in minimizing the cost function K are: the number of points sampled in each stage, the order in which the points are sampled (referred to as "ordering"), and the similarity criterion that is employed (e.g. normalized cross-correlation), and the similarity threshold for each stage. The threshold for the first stage is chosen to achieve the required probability of detection, while keeping the false alarms to a minimum. The threshold in the second stage is chosen to achieve the required probabilities of detection and false alarm for the entire process.

The measure of similarity between the input and reference that is chosen depends on the application in question. In digital image registration, the mean square error between the input scene points and template points is often used⁴. In pattern recognition applications, where the input scene may contain clutter, cross-correlation is often used as a similarity measure. In the following section we demonstrate how a photon-limited correlation signal is used to provide an estimate for the cross-correlation similarity measure.

6.4 Two-Stage Template Matching using Quantum-Limited Images

Photon-counting techniques are applied to two-stage template matching in the following manner. In the first stage, a reference function, or template is moved to each offset within the input scene. At each offset, the input is sampled using a small number of detected photoevents, $N_{1,p}$. The photon-limited cross-correlation $C(x,y)$ at each offset (x,y) is given by

$$C(x,y) = \sum_{j=1}^{N_{1,p}} R(x + x_j, y + y_j) \quad , \quad (6.2)$$

where (x_j, y_j) denote the coordinates of the detected photoevents. Equation (6.2) provides a measure of the similarity between the input and reference. The cross-correlation is compared with some pre-determined threshold $C_{T,1}$; if $C(x,y) > C_{T,1}$, that offset location is examined in the second stage using a sufficient number of detected photoevents to achieve the required probabilities of detection and false alarm for the particular application being considered. This process is repeated at each location in the input scene. It is important to note that nature provides the sampling statistics for the input scene. As mentioned earlier, the probability of detecting a photoevent at a given location is directly proportional to the classical intensity at the corresponding location in the input scene [see Eq. (2.7)]. This eliminates the need for the generation of random numbers, as is required in the standard digital implementation of the two-stage template matching technique. In the next subsection, the cost function for the photon-limited implementation of the two-stage template matching process is derived, which will allow optimum values for the the number of detected photoevents and correlation thresholds to be chosen for each stage of the search.

6.4.1 Derivation of the Quantum-Limited Cost Function

To optimize the photon-limited implementation of the two-stage template-matching algorithm, it is necessary to define a cost function similar to the one defined for the digital implementation. For the case of photon-counting, the cost function, denoted by κ , becomes

$$\kappa = N_{1,p}M + N_{2,p}\chi \quad , \quad (6.3)$$

where $N_{1,p}$ denotes the number of photon detected in the first stage at each offset, M denotes the the number of locations in the first stage, $N_{2,p}$ denotes the number of

detected photoevents in the second stage, and χ denotes the number of locations that exceeded the correlation threshold for the first stage, and must be examined in the second stage. In other words, the cost of the two stage process is defined in terms of the average number of photoevents detected. Typically, the number of locations that must be examined in the first stage, M , will be determined by the geometric dimensions of the input scene and reference function. For example, if input scene has dimension $m \times m$, and the reference function has dimension $k \times k$, then M will be given by $(m-k)^2$.

The number of locations that exceed the correlation threshold in the first stage, χ , will be some random function of the photon-limited correlation signals realized at each offset in the first stage. As a result, it is necessary to choose values for $N_{1,p}$ and $N_{2,p}$, as well as the values for the correlation thresholds used in the first and second stages, (denoted by $C_{T,1}$ and $C_{T,2}$ respectively), that will minimize the expected value of the cost function κ . In addition, the number of detected photoevents and correlation thresholds must be chosen such that the required probabilities of detection and false alarm are satisfied.

The expected value of the cost function κ , denoted by $\langle \kappa \rangle$, is given by

$$\langle \kappa \rangle = N_{1,p}M + N_{2,p} \langle \chi \rangle \quad . \quad (6.4)$$

The expected value of the number of offsets that will exceed the correlation threshold in the first stage is given by¹⁴

$$\langle \chi \rangle = \sum_{i=1}^M P(C_{i,1} > C_{T,1}) \quad , \quad (6.5)$$

where $P(C_{i,1} > C_{T,1})$ denotes the probability that the value of the photon-limited correlation signal realized in the first stage, at the i^{th} location C_i , will exceed the threshold $C_{T,1}$. The probability $P(C_{i,1} > C_{T,1})$ is given by

$$P(C_{i,1} > C_{T,1}) = \int_{C_{T,1}}^{\infty} P(C_{i,1}) dC_{i,1} \quad , \quad (6.6)$$

where $P(C_{i,1})$, denotes the probability density function for the photon-limited correlation signal in the first stage, at the i^{th} location in the input scene. If the reference function is real, and the number of detected photoevents is fixed, then $P(C_{i,1})$ is monovariate normal, as shown in Eq. (2.56). Using Eqs. (2.56) and (6.6), and substituting Eq. (6.5) into Eq. (6.4), it is possible to obtain an expression for the expected value of the cost function, $\langle \kappa \rangle$. The expected value of the cost function $\langle \kappa \rangle$ is given by

$$\langle \kappa \rangle = N_{1,p}M + N_{2,p} \sum_{i=1}^M \frac{1}{\sqrt{2\pi}\sigma_{i,1}} \int_{C_{T,1}}^{\infty} \exp \left\{ -\frac{(C_{i,1}(x_i, y_i) - \langle C_{i,1}(x_i, y_i) \rangle)^2}{2\sigma_{i,1}^2} \right\} \quad (6.7)$$

In Eq. (6.7) the mean values and standard deviations of the the photon-limited correlation signal that is realized in the first stage, at the i^{th} location in the input scene, are denoted by $\langle C_{i,1} \rangle$ and $\sigma_{i,1}$ respectively, and can be obtained using Eqs. (2.54) and (2.55). Equation (6.7) is the expression that one must minimize using the variables $C_{T,1}$, $C_{T,2}$, $N_{1,p}$ and $N_{2,p}$ subject to the restriction that the required probability of detection and false alarm be satisfied. Obtaining a solution for the values of the variables that will minimize the cost function κ would involve the simultaneous solution of a number of coupled integral equations, equal to the number of locations in the input scene; for general reference and input objects, this is analytically impossible.

Additionally, it is impractical to arrive at the answer by numerically solving the coupled integral equations. Alternatively, a more practical, iterative method is outlined in the following section.

6.4.2 Determination of Correlation Thresholds and Numbers of Detected Photoevents

The number of detected photoevents and correlation thresholds in each stage must be chosen to minimize the expected value of the cost function given in Eq. (6.7), and must provide the specified probabilities of detection and false alarm for the entire process. If it is desired to identify a particular object from within a given scene, the number of detected photoevents required to identify the object within specified probabilities of detection and false alarm can only be determined *exactly* by applying the method of hypothesis testing [described in Section 2.5] at *every* location in the input scene. This is due to the fact that the hypothesis testing technique requires knowledge of all of the objects from which the reference image must be chosen. Clearly, in an application such as the location of a reference image from within some unknown cluttered scene, it is not possible to satisfy this requirement. In this section, the approach that is taken is to place the reference object within a "typical" scene, or a scene that is most likely to be encountered in a given application. The number of detected photoevents required to identify the object from within that scene is then determined. It is important to note that if the object is placed in a different scene, the number of detected photoevents as determined from the "typical" scene may or may not be optimum; this must be tested on a case-by-case basis.

6.4.2.1 Stage Two

The number of detected photoevents that are utilized in the second stage of the two-stage template matching process, $N_{2,p}$, and the correlation threshold in the second stage, $C_{2,T}$, must be chosen to satisfy the requirements for the probability of detection and false alarm for the entire process. The over all probability of false alarm, denoted by P_F is given by

$$P_F = \sum_{i=1}^M P(C_{i,1} > C_{T,1}) P(C_{i,2} > C_{T,2}) \quad , \quad (6.8)$$

where $P(C_{i,1} > C_{T,1})$ denotes the probability that the value of the photon-limited correlation signals realized in the first stage, at the i^{th} location $C_{i,1}$, will exceed the threshold $C_{T,1}$. In Eq. (6.4), the index i is summed over all locations in the input that **do not** correspond to the reference object. Similarly, $P(C_{i,2} > C_{T,2})$ denotes the probability that the value of the photon-limited correlation signals realized in the second stage, at the i^{th} location $C_{i,2}$, will exceed the threshold $C_{T,2}$. The probability $P(C_{i,1} > C_{T,1})$ is given by Eq. (6.6). Similarly, the probability $P(C_{i,2} > C_{T,2})$ is given by

$$P(C_{i,2} > C_{T,2}) = \int_{C_{T,2}}^{\infty} P(C_{i,2}) dC_{i,2} \quad , \quad (6.9)$$

where $P(C_{i,2})$, denotes the probability density function for the photon-limited correlation signal in the second stage, at the i^{th} location in the input scene. If the reference function is real, and the number of detected photoevents is fixed, then $P(C_{i,1})$, and $P(C_{i,2})$ are obtained using Eq. (2.56). Hence,

$$P(C_{i,j}(x,y)) = \frac{1}{\sqrt{2\pi}\sigma_{i,j}} \exp - \left\{ \frac{C_{i,j}(x,y) - \langle C_{i,j}(x,y) \rangle}{2\sigma_{i,j}^2} \right\} \quad . \quad (6.10)$$

The mean value and variance in Eq. (6.7) are obtained using Eqs. (2.54) and (2.55) respectively, i.e.,

$$\langle C_{i,j}(x,y) \rangle = N \iint_A p_{i,j}(x',y') R(x+x',y+y') dx' dy' \quad , \quad (6.11)$$

and

$$\sigma_{i,j}^2 = N \iint_A p_{i,j}(x',y') R^2(x+x',y+y') dx' dy' - \frac{\langle C_{i,j} \rangle^2}{N} \quad , \quad (6.12)$$

where the number of detected photoevents N is given by $N_{1,p}$ or $N_{2,p}$ as appropriate. In Eqs. (6.10)-(6.12), the subscript "i" refers to the offset location in the input scene, and the subscript "j" refers to the stage in the search (e.g. 1 or 2).

Equation (6.8) shows that the over-all probability of false alarm will depend on the correlation thresholds and the number of detected photoevents in both stages. As a result, it is not straight forward to apply the hypothesis testing approach to determine $C_{T,2}$ and $N_{2,p}$. However, an approach that is effective is to choose $C_{T,2}$ and $N_{2,p}$ as if *all* of the locations examined in the first stage must be examined in the second stage. In this case, the overall probability of false alarm is considered to be given by

$$P_F = \sum_{i=1}^M P(C_{i,2} > C_{T,2}) \quad , \quad (6.13)$$

where once again, the summation is performed over all locations in the input that do not correspond to the reference object. Here, the first estimates for the correlation threshold and number of detected photoevents in the second stage are chosen independent of $C_{T,1}$ and $N_{2,1}$, by applying the hypothesis testing technique described in Section 2.5. Initial estimates are chosen for $N_{2,p}$ and $C_{T,2}$ such that the probability of detection is satisfied when the input and reference objects are the same (i.e. the template matches the input). The overall probability of false alarm, P_F , is computed for

a typical input scene using Eq. (6.13). If the requirement for P_F is not satisfied, the number of detected photoevents $N_{2,p}$ is increased, and the above process repeated until the requirement for P_F is satisfied. When this process is complete, one has obtained the first estimate for $N_{2,p}$ and $C_{T,2}$, and can then proceed to obtain estimates for $N_{1,p}$ and $C_{T,1}$ as outlined in the next section.

6.4.2.2 Stage One

The number of detected photoevents and correlation threshold in the first stage, denoted by $N_{1,p}$ and $C_{T,1}$ respectively, must be chosen such that the required probability of detection is maintained, while minimizing the expected value of the cost function given in Eq. (6.7). The method for determining $N_{1,p}$ and $C_{T,1}$ that the author has found to be most effective is an iterative one. The cost function in Eq. (6.7) is plotted over a small range of detected photoevents; it has been the author's experience that a typical starting value for $N_{1,p}$ can be obtained by using $N_{2,p}/15$, where $N_{2,p}$ is the estimate obtained for the number of detected photoevents in the second stage. In addition, the range of detected photoevents over which the cost function must be plotted is typically around 50 photoevents. One must examine the resulting cost function, and determine whether a minimum lies in the selected range of detected photoevents, and make adjustments if necessary. The correlation threshold $C_{T,1}$ must be determined for each value of $N_{1,p}$. The correlation threshold $C_{T,1}$ is selected such that probability of detection for the overall process is satisfied when the input and reference objects are the same. This probability of detection is computed using Eq. (6.6). The probability density function $P(C_{i,1})$ is given by Eq. (6.10), using the mean value and standard deviation given in Eqs. (6.11) and (6.12), respectively, when the input and reference objects are the same. One can then examine the plot of Eq. (6.7), and obtain the value for $N_{1,p}$ for the given choice of $N_{2,p}$. One can then repeat the process for obtaining

the estimate for $N_{2,p}$, using Eq. (6.8) to compute the probability of false alarm. If necessary, the entire process can be repeated several times, but it has been the author's experience that repeating this process only decreases the cost function by a few percent. A summary of the method for determining the correlation threshold and number of detected photoevents is given on the following page.

Determination of $N_{1,p}$, $N_{2,p}$, $C_{T,1}$, and $C_{T,2}$: Summary

Stage Two

- 1.) Make initial estimate for $N_{2,p}$; choose $C_{T,2}$ as large as possible, while satisfying the requirement for the probability of detection for the given reference object (template). This is performed using Eqs. (6.9), along with Eqs. (6.10)-(6.12).
- 2.) Compute the overall probability of false alarm P_F for the reference object in a given scene using Eq. (6.13).
- 3.) If the specified probability of detection is satisfied, decrease the value for $N_{2,p}$ and repeat steps (1) and (2).
- 4.) If the specified probability of detection is not satisfied, increase the value for $N_{2,p}$ and repeat steps (1) and (2).
- 5.) A high-low search is then performed for the value of $N_{2,p}$ that most nearly satisfies the requirements for the probability of detection and false alarm.

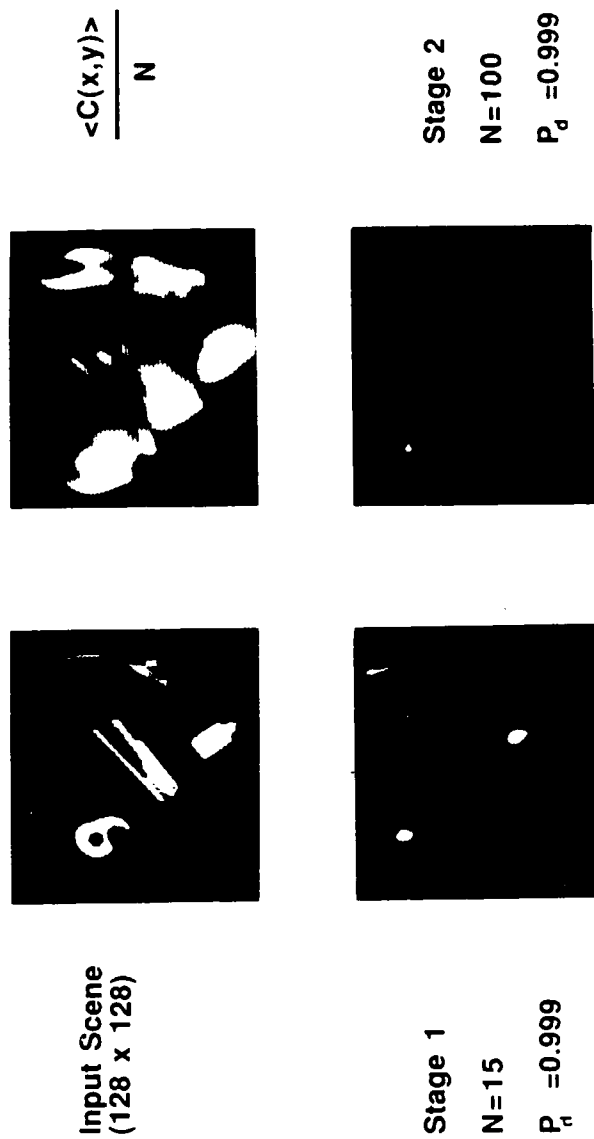
Stage One

- 6.) Use $N_{2,p}/15$ as the start of the range of values for $N_{1,p}$ over which the cost function is to be plotted. A range of approximately 50 is usually sufficient.
- 7.) For each $N_{1,p}$, choose a corresponding value for $C_{T,2}$ that is as large as possible, while satisfying the requirement for the probability of detection for the given reference object. This is done using Eq. (6.6), along with Eqs. (6.10)-(6.12).
- 8.) Plot the cost function in Eq. (6.7) over the given range of detected photoevents.
- 9.) If a minimum is observed, use the corresponding value for $N_{1,p}$ in steps (1)-(3) to update the value for $N_{2,p}$.
- 10.) If a minimum is not observed, a high-low search can be performed for the range of values of $N_{1,p}$ that contain the minimum of the cost function. Once this range is obtained, steps (7)-(9) are then repeated.

6.5 Computer Simulations

The following computer simulations demonstrate the theoretical predictions for the performance of two-stage template matching using quantum-limited images in several different environments. Figure 6.2 demonstrates a controlled, uncluttered environment, or an environment in which the objects in question have been segmented from the background via some pre-processing. In Fig 6.2, it was desired to locate and identify the tape dispenser, shown in the upper left. The input scene was digitized to a resolution of 128x128 pixels, and the template was 32x32 pixels. The upper right denotes the expected value of the photon-limited correlation signal divided by the number of detected photoevents N . This is analogous to the classical-intensity cross-correlation function. The procedure outlined in subsection 6.4 was used to choose the number of detected photoevents and correlation thresholds in each stage. The lower left and lower right of Fig. 6.2 denotes the probability that the correlation threshold will be exceeded at each offset for the number of detected photoevents indicated. Each picture has been normalized such that a probability of 1 was mapped into a grey level of 255, and a probability of 0 was mapped into 0.

Reference Object: Tape Dispenser (32x32)



Probability that the similarity criterion
will be satisfied at each window offset.

Figure 6.2 Two-stage template matching in a scene with segmented objects.

In the first stage, the correlation threshold $C_{T,1}$ was chosen such that a probability of detection of 0.999 was achieved; the average number of points that exceeded the threshold was 80. In the second stage, $C_{T,2}$ was chosen such that P_d remained 0.999, and the probability of false alarm was reduced to 0.001. The number of offsets examined in the first stage is $(128-32)^2=9216$. In the second stage, 80 points on average need to be examined. Presently available, position-sensitive, photon-counting detection systems operate at rates up to 100 KHz; hence the entire two-stage search process will take approximately 1.4 seconds.

In the example shown in Fig. 6.3, a template matching application involving satellite images was investigated. The 32x32 reference image shown in Fig. 6.3 was arbitrarily chosen from the 128x128 input scene. In Fig. 6.4, the cost function in Eq. (6.7) is plotted using the given input and reference functions, for $N_{2,p}=1000$. Figure 6.4 was used to select the given value of $N_{1,p}$. Note that while the exact minimum in the cost function occurred with a value of 96 for $N_{1,p}$, this may not be exactly the optimum number when a slightly different scene is used; hence for convenience, 100 detected photoevents were used in the computer simulation. In the lower left of Fig. 6.3, the probability that $C(x,y)$ will exceed the value chosen for $C_{T,1}$ is plotted at each offset location. Again, a probability of one was mapped into a gray level of 255. One hundred detected photoevents were used in the first stage, with $C_{T,1}$ chosen such that P_d was 0.95. On average, 3700 points need to be examined during the second stage. In the second stage, offset locations exceeding the threshold were examined using 1000 detected photoevents. The probability that each offset location will exceed the value chosen for $C_{T,2}$ is shown in the lower right. The threshold $C_{T,2}$ was chosen such that P_d in the second stage remained 0.95, with a P_{fa} of 0.001. Hence, the entire search

process would take approximately 45 seconds, based on detection rates of 100 KHz. Recent work^{15,16} suggests that detection rates on the order of 1 MHz. may be possible, in which case the detection times above could be reduced by an order of magnitude.

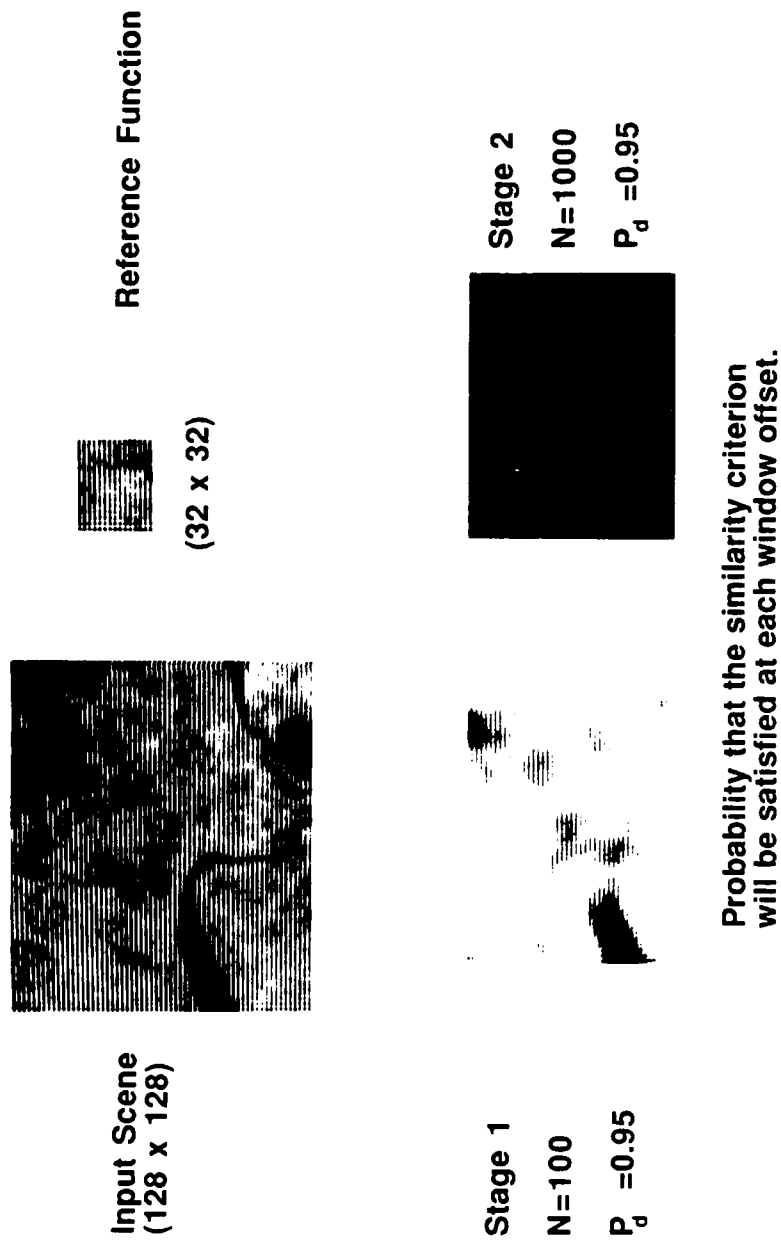


Figure 6.3 Two-stage template matching in a satellite image.

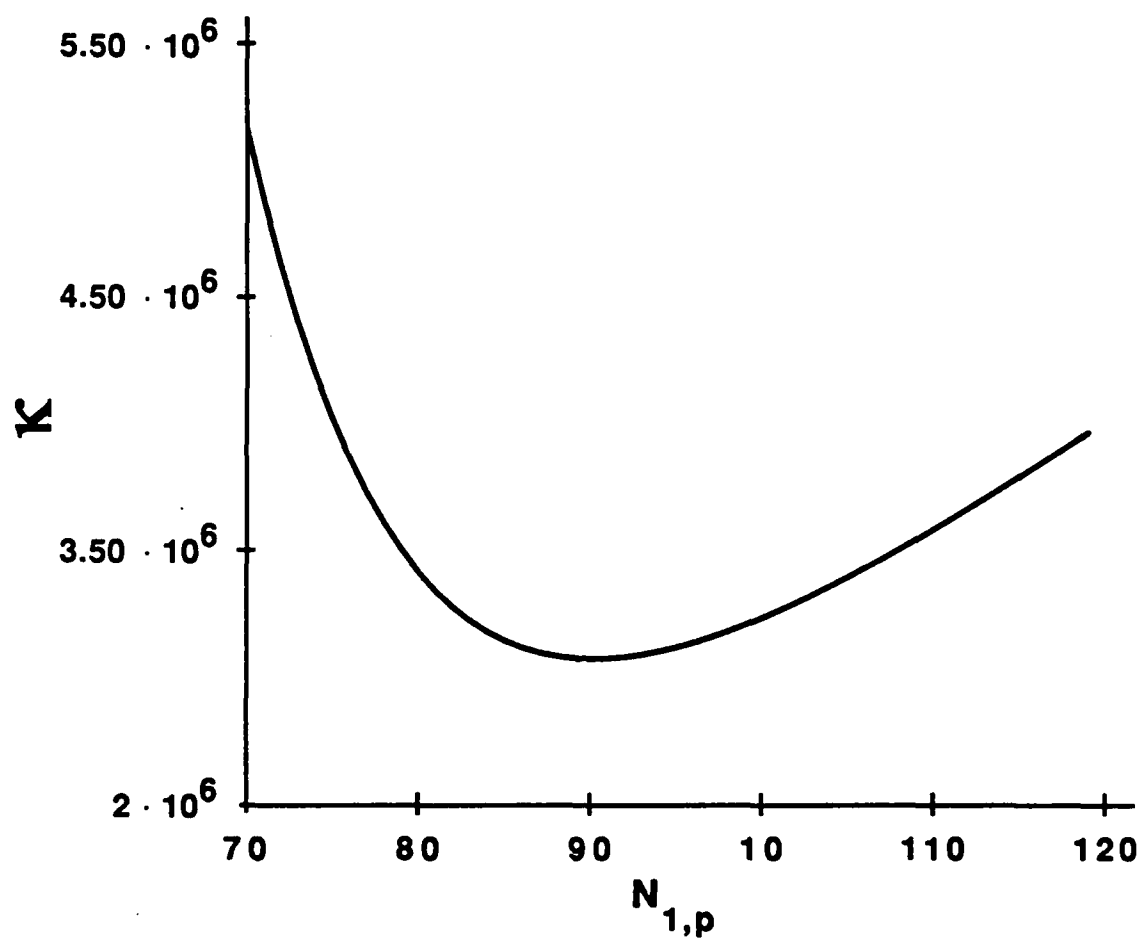


Figure 6.4 Cost function for the input and reference functions shown in Fig. 6.3.
The number of detected photoevents used in the second stage was 1000.

Figure 6.5 demonstrates an attempt at two-stage template matching using a reference function with insufficient resolution (20x40 pixels) to accurately detect the reference objects from the cluttered 512x512 input scene. In this case, it is not possible to eliminate the false alarms, no matter how many detected photoevents are used in stage two.

In Figure 6.6, the resolution of the reference image was increased to 32x50 pixels. In the first stage of the search, 50 photoevents are collected at each coordinate of the 512x512 pixel input scene, and the correlation threshold is set to give a probability of detection equal to 0.99. There are 4000 locations identified as likely locations for the targets (most of the locations are located on the targets). In stage two, 300 detected photoevents were collected at each of the positions identified by stage one. The probability that the correlation output will exceed the threshold at each location in stage two is shown in the lower right of Fig. 6.6. Note that all trucks were identified, with one false alarm. The number of detected photoevents utilized in the first stage of the search were chosen using the cost function shown in Fig. 6.7. The total number of detected photoevents employed in the two-stage search is 1.4×10^7 .

In Fig. 6.7, the minimum of the cost function occurred at a value of $N_{1,p}=32$; as before, because that particular value of $N_{1,p}$ may not be optimum when the target object is placed in a different scene, the value of 50 was chosen for convenience. It is important to note that the cost function only varies by about 10% over the range of values of $N_{1,p}$ between 25 and 50. This makes it possible to quickly estimate a reasonable value of $N_{1,p}$ without plotting the entire function.

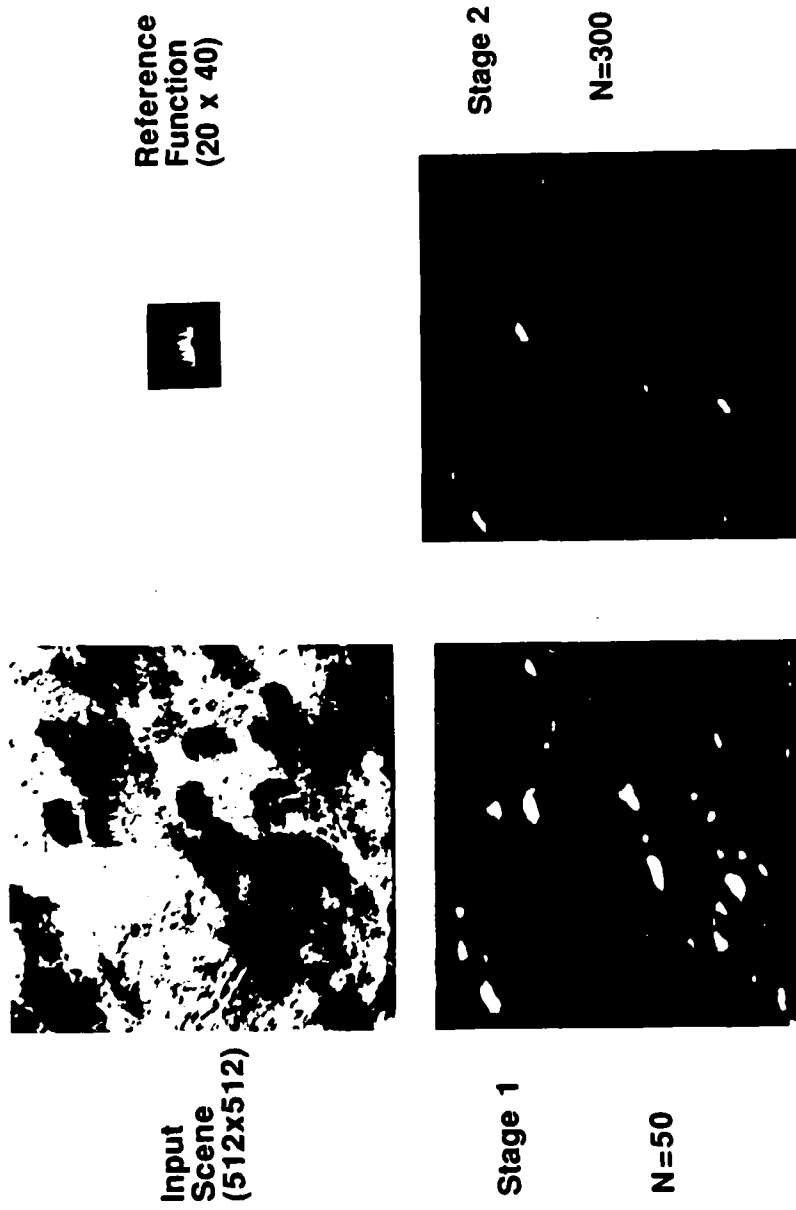


Figure 6.5 Attempt at two-stage template matching in an aerial photograph using a reference function with insufficient resolution.

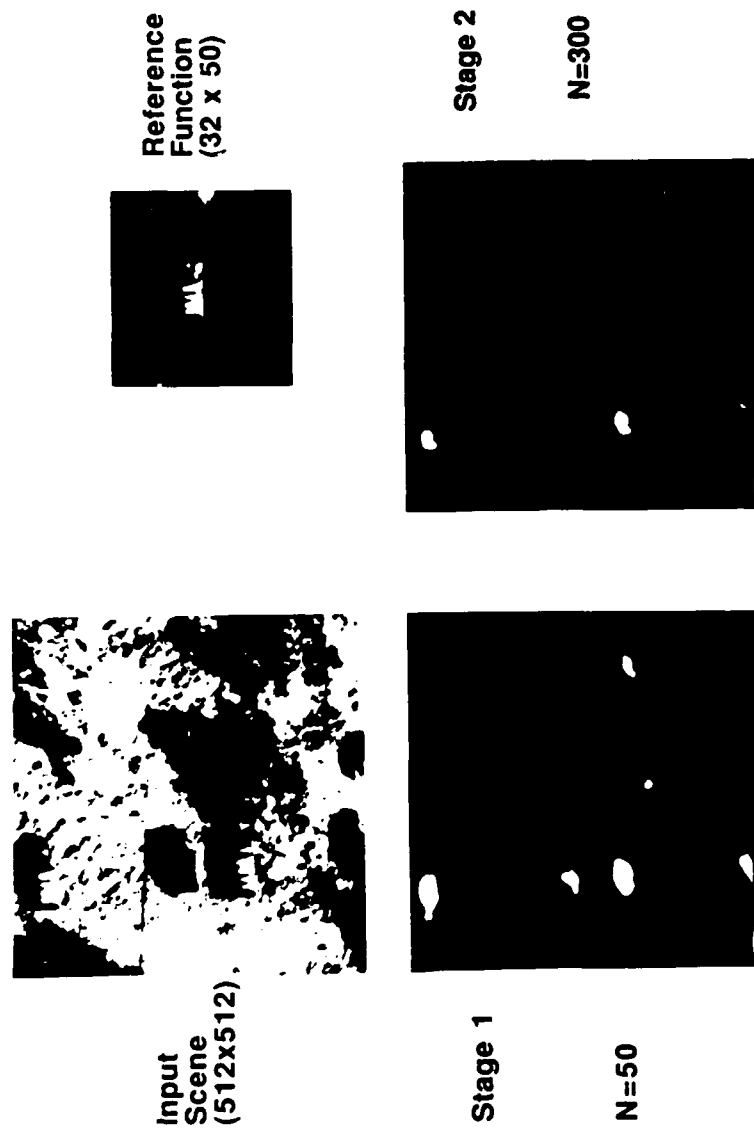


Figure 6.6 Two-stage template matching in an aerial photograph using a reference function with increased resolution.

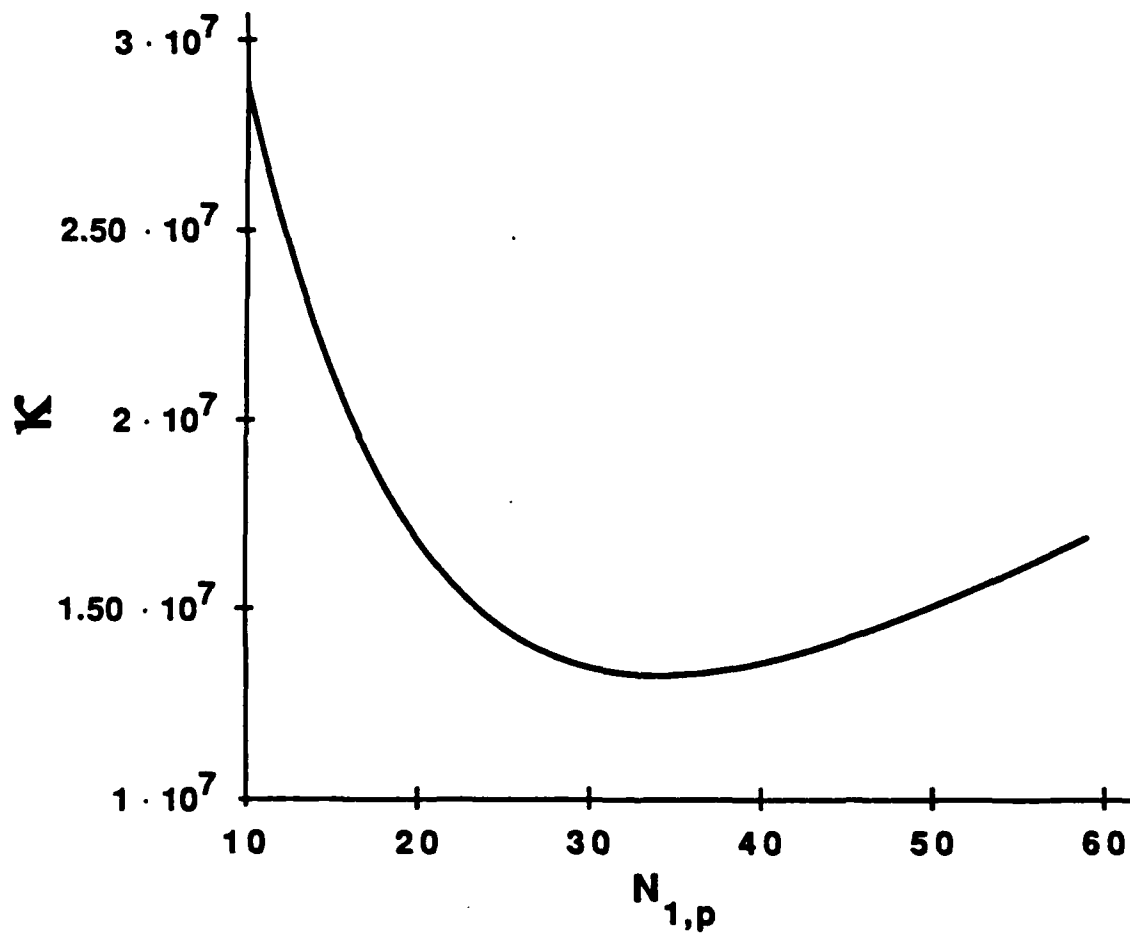


Figure 6.7 Cost function for the input and reference functions shown in Fig. 6.6. The number of detected photoevents used in the second stage was 300.

6.6 Two-Stage Invariant Filtering using Circular-Harmonic Functions

The same approach that is described in Section 6.5 for real reference functions can be used for complex filter functions, such as the circular-harmonic filter described in Chapter 4. Equation (4.24), is used as the probability density function for the normalized correlation signal from which the necessary probabilities of detection and false alarm are calculated. Figure 6.8 demonstrates the performance of the rotation-invariant circular-harmonic filter described in Chapter 4, with the correlation signal normalized at each offset using the method described in Section 4.5.2. In Fig. 6.8, the input scene (top left, 512x512 pixels) is the same input scene used in Fig. 4.8. The reference object was once again the top truck in the center of the input scene, from which the circular-harmonic filter (top right) was computed, about its proper center. Theoretical predictions for the probability that the correlation signal realized using the indicated number of detected photoevents at each location will exceed the correlation threshold are shown in the bottom of Fig. 6.8. The over all probability of detection for the four single trucks that was achieved was 0.999, while there was one continuous area that produced false alarms. The "double truck" was not identified.

6.7 Discussion

For input images such as those in Figs. 6.6 and 6.8, searching each offset location using different detected photoevents can require the detection of up to 1.4×10^7 photoevents (see Fig. 6.7). Clearly, even operating at detection rates of 1 MHz., this method may not be practical. Fortunately, by examining the properties of the output correlation function, the total search time can be further decreased. Note that in the

output from stage one in Figs. 6.6, and 6.8, the output correlation functions are rather wide. This is due to the relatively low space-bandwidth product of the target images. As a result, it is possible to search every third or fourth location in stage one, rather than performing an exhaustive search. If one were to search every fourth pixel, and employed a detection system that operated at 1 MHz., the entire two-stage search of the input scene in Fig. 6.6 would take approximately 0.8 seconds if 32 detected photoevents were used in the first stage. Clearly, this is a more encouraging performance. It is important to note that the same objective might sometimes be obtained by reducing the resolution of the input scene in the first stage³; however, the results shown in Fig. 6.5 indicate that approach may not always be as effective. (Note that in Fig. 6.5, a large number of locations exceeded the threshold in the first stage.)

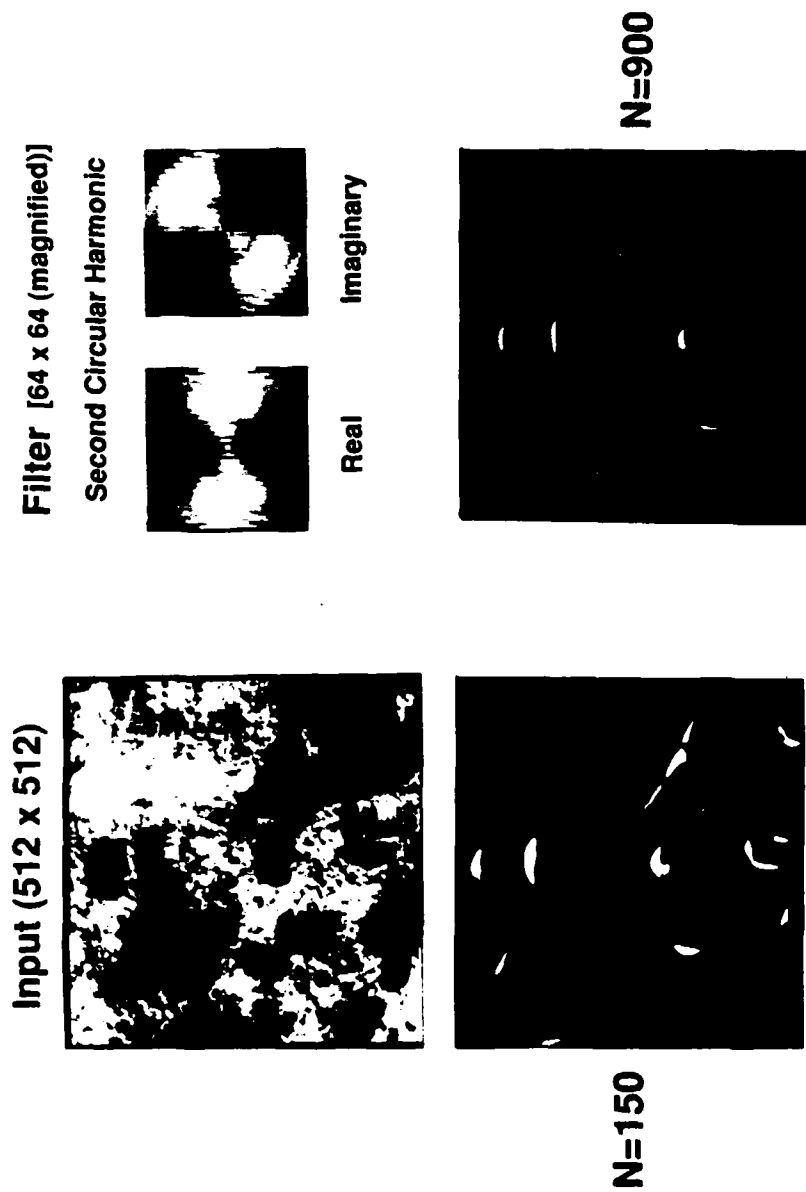


Figure 6.8 Two-stage invariant filtering. Upper left, input scene; upper right, reference circular-harmonic filter. Lower left, probability that the photon-limited correlation realized using 150 photoevents will exceed the similarity criterion at each offset. Lower right: same as lower left, using 900 detected photoevents.

6.8 Summary

In this Chapter, we have investigated two-stage template matching using quantum-limited images. In Section 6.2, a brief review of the two-stage template matching method for image recognition is given. The theoretical basis for template matching using quantum-limited images is given in Section 6.3. A summary of the procedure for choosing the number of detected photoevents and correlation thresholds in each stage is provided in Section 6.4.2.2. In Section 6.5, theoretical predictions are given for the performance of two-stage template matching using quantum-limited images in: a segmented scene, template matching in a satellite image, and automatic target recognition (see Figs. 6.2-6.7). Finally, two-stage template matching using circular-harmonic filters is considered, with encouraging results (see Fig. 6.8).

References for Chapter 6

1. J. S. Bomba, 'Alpha-numeric character recognition using local operations,' Fall Joint Comput. Conf., 218-224 (1959).
2. D. Barnea and H. Silverman, 'A Class of Algorithms for Fast Digital Image Registration,' IEEE Trans. Comput. C-21, 179-186 (1972).
3. A. Rosenfeld and G. Vanderburg, 'Coarse-Fine Template Matching,' IEEE Trans. Sys. Man, Cyber., 104-107 (1977).
4. G. Vanderburg and A. Rosenfeld, 'Two-Stage Template Matching,' IEEE Trans. Comput. C-26, 384-393 (1977).
5. A. Goshtasby, S. Gage, and J. Bartholic, 'A Two-Stage Cross Correlation Approach to Template Matching,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-6 374-378 (1984).
6. C. J. Oliver, 'An analysis of template matching in image registration,' Opt. Acta 31, 233-348 (1984).
7. X. Li and R. C. Dubes, 'The first stage in two-stage template matching,' IEEE Trans. Patt. Anal. Mach. Intell. PAMI-7, 700-707 (1985).
8. J.L. Turney, T.N. Mudge and R.A. Volz, 'Recognizing partially occluded parts,' IEEE Trans. Patt. Anal. Mach. Intell. PAMI-7, 410-421 (1985).
9. Z. X. Fang, L. Xiao, and L.M. Ni, 'Parallel algorithms for image template matching on hypercube SIMD computers,' IEEE Trans. Patt. Anal. Mach. Intel. PAMI-9, 835-841 (1987).
10. C.I. Soo, J.T. Walker, C.F. Chong, D.H. Damerson and R.F. Pease, 'Template set approach to defect detection and classification for VLSI patterns,' OSA technical Digest Series 12, 18-24 (1987).
11. D. Casasent and R. Krishnapuram, 'Hough space transformations for discrimination and distortion estimation,' Comput. Vis., Graph. Imag. Proc. 38, 299-316 (1987).
12. S. Ranka and S. Sahni, 'Image template matching in MIMD hypercube multicomputers,' Proc. 1988 Int. Conf. Parall. Process., Penn. State Univ., 15-19 (1988).
13. B. Bhanu, 'Automatic target recognition: state of the art survey,' IEEE Trans. Aerospace Elec. Sys. AES-22, 364-379 (1986)

14. T. A. Isberg and G. M. Morris, 'Two-Stage Template Matching using quantum-limited images,' *Proc. SPIE* **976**, 160-167 (1988).
15. S. E. Sobattka and M. B. Williams, 'Delay line readout of microchannel plates,' *IEEE Trns. Nucl. Sci.* **35**, 348-351 (1988).
16. J. A. Harder, 'A fast time-to-digital converter for position-sensitive radiation detectors with delay line readouts,' *Nucl Instrum. Meth. Phys. Res.* **a265**, 500-510 (1988).

Chapter 7 Concluding Remarks

In this Thesis, it is demonstrated that correlation-based methods for image recognition can be effectively implemented using a position-sensitive, photon-counting detection system. In Chapter 2, a theoretical formalism is developed that describes the behavior of a *correlation signal realized by cross correlating* a quantum-limited input scene with a reference function stored in computer memory. The theory describes the behavior when the reference function is in general complex, and the correlation signal is realized using either a fixed, or random number of detected photoevents. The method of hypothesis testing is described as it applies to the selection of the number of detected photoevents required to recognize a given image based on the photon-limited correlation signal. It is important to note that the theoretical formalism presented in this Thesis can be used to predict the effectiveness of **any** reference function, and not just the ones utilized in this Thesis.

In Chapter 3, image correlation at low light levels is investigated. Quantum-limited input images are cross-correlated with a digitized version of a classical-intensity reference image that is stored in computer memory; the resulting correlation signal corresponds to that of a conventional matched filter. The theory presented in Chapter 2 for the case of a real reference function and a fixed number of detected photoevents is verified experimentally, where excellent agreement between theory and experiment was observed.

It was theoretically predicted, and experimentally verified that as few as 1000 detected photoevents provide enough information to estimate the cross-correlation signal with enough accuracy to discriminate among a set of engraved portrait images. The experimental system employed can detect photoevents at rates up to 100 kHz. As a

result, the time to detect, process, and make a recognition decision regarding the input can be as little as 10 milliseconds.

In the experiments performed in Chapter 3, dark noise in the detection system was measured experimentally, and shown to be insignificant. However, a theoretical formalism was presented that allows one to predict the effect of additive noise on the recognition performance of the system. In addition, a method for reducing the effect of the additive noise was suggested, and computer simulations were performed that demonstrated the effectiveness of this method.

In Chapter 4, the performance of the rotation-invariant circular-harmonic filter was investigated under quantum-limited conditions. It was theoretically predicted, and experimentally observed, that as few as 3000 detected photoevents are all that are needed to accurately recognize a segmented object, and determine its orientation. Excellent agreement was observed between theory and experiment.

Circular-harmonic filters have been known to perform poorly in applications that require detection of objects within a cluttered environment. This is due to the fact that is computationally impractical to normalize the output correlation signal properly using conventional methods. In Chapter 4, a new normalization is suggested that, while not optimum, may provide improved performance when the circular-harmonic filter is used in a cluttered environment. This new normalization method can be performed in real time using photon-counting techniques. This new method of normalization is employed in an application to automatic target recognition in an aerial photograph, with encouraging results. In the future, extensive testing needs to be performed to fully determine the usefulness of this method for normalization.

In Chapter 5, the estimation of moment invariants for pattern recognition is investigated. It is demonstrated that the strong, or largest moments of input images

can be accurately estimated using photon-counting techniques. Fortunately, these are the moments that are typically useful for pattern recognition applications. The number of detected photoevents required to estimate the moments of segmented input objects are theoretically predicted, and experimentally verified. Using 5000 detected photoevents, it is possible to estimate moment invariants that can be used to discriminate among a set of engraved-portrait images that are input at any orientation, and at scales that vary up to a factor of two. This was the first real-time method reported for computing moment-invariants with sufficient accuracy for pattern recognition applications with complex objects. Unfortunately, moment-invariants are not useful for recognizing objects that are not segmented from the background, and are not usually utilized in cluttered environments.

Finally, in Chapter 6, the implementation of a two-stage template matching algorithm using photon-counting techniques is considered. In this photon-counting method for image recognition, the correlations are implemented in an image plane; as a result, one does not enjoy the position-invariance that exists when the correlations are implemented optically, or digitally using fast Fourier transforms (FFT's). As a result, if one desires to recognize objects from within a cluttered environment, one must search the input scene by moving a reference window to different locations within the input scene, and realize a photon-limited correlation signal at each location. This technique is known as template matching, and has a long history in the pattern recognition literature.

Here, a two-stage template matching technique is investigated. In the first stage, the input scene is quickly searched by detecting a few photoevents at each location in the input, and comparing the resulting photon-limited correlation signal to some pre-determined threshold. If the correlation signal at a particular location exceeds the threshold, that location is determined to be a likely location for the target, and is

examined fully in the second stage of the search.

In Chapter 6, this two-stage template matching approach is applied to: an input scene consisting of segmented objects, scene matching in a satellite image, and automatic target recognition in a cluttered environment. In realistic scenes, it is demonstrated that only a few tens of detected photoevents may be required in the first stage, and only a few hundred photoevents may be required in the second stage. However, if a 512×512 scene is searched at every possible location, it may be necessary to detect as many as 1.4×10^7 photoevents. Using the detection system employed in this Thesis, which can detect photoevents at rates up to 1×10^5 per second, this method may prove impractical in some applications.

Recently, position-sensitive photon-counting detectors have been reported to operate at rates up to 1 MHz. In addition, in the images that were considered here, the output correlation functions were relatively broad; as a result, it may be possible to sample every third or fourth pixel in the first stage of the search. The new technology, combined with searching every third or fourth pixel would decrease the search time dramatically. It is important to note two things regarding this approach. First, at detection rates approaching 1 MHz., it is important to take into consideration the effects of dead time on the output correlation signal¹. Secondly, searching every third or fourth pixel in the first stage is not the same as reducing the resolution in the first stage of the search (i.e., by digitizing the photoevent coordinates to 7 bits instead of 8). This was demonstrated in Chapter 6, where a reference object digitized to a resolution of 20×40 pixels could not be accurately identified from a cluttered scene, but could be identified when it was digitized to a resolution of 32×50 . In the future, the use of classification filters should be employed in this two-stage template matching process, which may allow the target images to be recognized under more cluttered conditions.

¹Doo-Jin Cho, *Dead time effects in photon-counting*, Ph. D. Thesis, University of Rochester, (1989) Chap. 4.

Appendix A: Statistics of the photon-limited correlation signal

This appendix is divided into two sections. In Section A.1, theoretical expressions for the statistical quantities involving the photon-limited correlation signal are given. Many of the equations presented in this appendix are similar to those presented in Chapter 2, but vary slightly because of the way the photon-limited correlation signal is defined here [compare Eqs. (2.9) and (A3)] for the particular application of moment estimation. Section A.2 provides a description of the statistics for the photon-limited estimate of the distance in feature space between an input image and a reference image. Finally, Section A.3 contains a description of the method used to determine the number of detected photoevents required to specify the centroid of the input to within a given error.

A.1 Statistics of the modulus of the correlation signal

A photon-limited input image $f^\dagger(r, \theta)$ can be represented as a two-dimensional collection of Dirac delta functions, i.e.,

$$f^\dagger(r, \theta) = \sum_{i=1}^N \delta(r - r_i, \theta - \theta_i) \quad (A1)$$

in which (r_i, θ_i) denotes the spatial coordinates of the i th detected photoevent and N is the total number of detected photoevents. Here, we consider the case where the number of detected photoevents N is fixed, which makes the coordinates of the detected photoevents (r_i, θ_i) the only random variables. The probability density function for the detected photoevent coordinates is directly proportional to the classical intensity of the corresponding location in the input scene^{1,2}, i.e.,

$$P(r_i, \theta_i) = \frac{f(r_i, \theta_i)}{\int_0^{2\pi} \int_0^\infty f(r, \theta) r dr d\theta} \quad (A2)$$

The cross correlation $C(r, \theta)$ between a photon-limited input image $f^\dagger(r, \theta)$ and a complex reference function $R(r, \theta)$ is given by (c.f. Eq.(11))

$$C(r, \theta) = \frac{1}{N} \sum_{i=1}^N R(r + r_i, \theta + \theta_i) \quad (A3)$$

In general, the reference function is complex, making the correlation signal $C(r, \theta)$ complex. For convenience, let

$$C(r, \theta) = C'(r, \theta) + iC''(r, \theta) \quad (A4)$$

where C' and C'' represent the real and imaginary parts of the correlation signal, respectively. For sufficiently large N , the joint probability density function for C' and C'' is approximately a bivariate normal density, given by

$$P(C', C'') = \frac{1}{2\pi\sigma'\sigma''(1-\rho^2)^{\frac{1}{2}}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(C' - \langle C' \rangle)^2}{\sigma'^2} - 2\rho \frac{(C' - \langle C' \rangle)(C'' - \langle C'' \rangle)}{\sigma'\sigma''} + \frac{(C'' - \langle C'' \rangle)^2}{\sigma''^2} \right] \right\}, \quad (A5)$$

where the mean value of the real and imaginary parts of the correlation signal, $\langle C' \rangle$ and $\langle C'' \rangle$ are

$$\langle C'(r, \theta) \rangle = \frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') \operatorname{Re}\{R(r+r', \theta+\theta')\} r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'}, \quad (A6)$$

and

$$\langle C''(r, \theta) \rangle = \frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') \operatorname{Im}\{R(r+r', \theta+\theta')\} r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'}, \quad (A7)$$

respectively. Note that the expression for the expected values are independent of the number of detected photoevents N , in contrast with the expressions for the expected values given in Eqs.(2.43) and (2.44).

The variances of C' and C'' are denoted by σ'^2 and σ''^2 , respectively. Using the same method that was described in Section 2.3.2, one finds

$$\sigma'^2 = \frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') [\operatorname{Re}\{R(r+r', \theta+\theta')\}]^2 r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'} - \frac{\langle C' \rangle^2}{N}, \quad (\text{A8})$$

and

$$\sigma''^2 = \frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') [\operatorname{Im}\{R(r+r', \theta+\theta')\}]^2 r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'} - \frac{\langle C'' \rangle^2}{N}. \quad (\text{A9})$$

respectively. The correlation coefficient r defined as

$$\rho = \frac{\langle C' C'' \rangle - \langle C' \rangle \langle C'' \rangle}{\sigma' \sigma''}, \quad (\text{A10})$$

is given by

$$\rho = \frac{1}{N \sigma' \sigma''} \left[\frac{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') \operatorname{Re}\{R(r+r', \theta+\theta')\} [\operatorname{Im}\{R(r+r', \theta+\theta')\}] r' dr' d\theta'}{\int_0^{2\pi} \int_0^{\infty} f(r', \theta') r' dr' d\theta'} \right. \\ \left. - \langle C' \rangle \langle C'' \rangle \right]. \quad (\text{A11})$$

In Eqs (A6)-(A11), $\operatorname{Re}\{\dots\}$ and $\operatorname{Im}\{\dots\}$ denote the real and imaginary parts of the reference function, respectively.

In the computation of the moment invariants, we require the marginal density function for the modulus of the correlation signal $|C(r, \theta)|$. To find the marginal density, we make the following change of variables:

$$C' = |C| \cos \gamma \quad (\text{A12.a})$$

$$C'' = |C| \sin \gamma \quad (\text{A12.b})$$

$$|C| = (C'^2 + C''^2)^{\frac{1}{2}} \quad (\text{A12.c})$$

Using the appropriate Jacobian for the transformation, substituting Eqs (A12a-c) into Eq.(A5), and integrating over γ , we obtain the following expression for the marginal density function $P(|C|)$:

$$P(|C|) = \frac{|C|}{4\pi(1-\rho^2)^{\frac{1}{2}}\sigma'\sigma''} \int_0^{2\pi} d\gamma \left\{ \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{|C|^2 \cos^2(\gamma) - 2|C| \langle C' \rangle \cos(\gamma) + \langle C' \rangle^2}{\sigma''^2} \right. \right. \right. \\ \left. \left. - 2\rho \frac{|C|^2 \cos(\gamma) \sin(\gamma) - |C| [\langle C'' \rangle \cos(\gamma) + \langle C' \rangle \sin(\gamma)] + \langle C' \rangle \langle C'' \rangle}{\sigma' \sigma''} \right. \right. \\ \left. \left. + \frac{|C|^2 \sin^2(\gamma) - 2 \langle C'' \rangle |C| \sin(\gamma) + \langle C'' \rangle^2}{\sigma''^2} \right] \right\} \right\} \quad (\text{A.13})$$

For the general form of $P(|C|)$ in Eq.(A13), we have been unable to obtain an analytic solution for the integral. However, in the particular case of the moment invariant computation, Eq.(A13) simplifies considerably. Consider the case when the normalized correlation coefficient ρ is small, and $\sigma'^2 = \sigma''^2 \equiv \sigma^2$ (These are experimental observations that we have observed to be true without exception, to date). Hence, is possible to neglect the middle term with respect to the other terms in the argument of the exponential in Eq.(A13). Making these approximations, and rearranging terms, we have

$$P(|C|) \equiv \frac{|C|}{2\pi(1-\rho^2)^{\frac{1}{2}}\sigma^2} \exp\left\{-\frac{|C|^2 + \langle C' \rangle^2 + \langle C'' \rangle^2}{2(1-\rho^2)\sigma^2}\right\} \times \int_0^{2\pi} d\gamma \exp\left\{\frac{|C|}{\sigma^2(1-\rho^2)} [\langle C' \rangle \cos \gamma + \langle C'' \rangle \sin \gamma]\right\} \quad (A14)$$

The integral in Eq.(A14) can be evaluated using Ref.3. Evaluating the integral in Eq. (A14) gives

$$P(|C|) \equiv \frac{|C|}{2\pi(1-\rho^2)^{\frac{1}{2}}\sigma^2} \exp\left\{-\frac{|C|^2 + \langle C' \rangle^2 + \langle C'' \rangle^2}{2(1-\rho^2)\sigma^2}\right\} I_0\left\{\frac{|C|(\langle C' \rangle^2 + \langle C'' \rangle^2)^{\frac{1}{2}}}{\sigma^2(1-\rho^2)}\right\} \quad (A15)$$

where $I_0\{\dots\}$ is the 0th - order modified Bessel function⁴. Hence, when the reference function in Eq.(13) is used, and $s=2$, then $P(|C|)$ represents the density function for the estimate of the invariant moment $\Phi_{2,m}$. Equation (A15) can be used to compute the mean value and variance for the estimate of $\Phi_{2,m}$ when Eq.(5.17) is used as the reference function, with $s=2$.

The description of the statistics of the estimate for $\Phi_{1,m}$ and $\Phi_{3,m}$ requires the computation of an additional PDF. As described in Section 5.3.2, the estimate for $\Phi_{1,m}$ is obtained by taking the ratio of the modulus of two correlation signals. [The correlation signals are obtained using the reference functions in Eqs. (5.24) and (5.25)]. Hence, we require the PDF for the random variable defined by

$$|C| = \frac{|C_1|}{|C_2|} \quad (A16)$$

where $|C_1|$ and $|C_2|$ are the moduli of the correlation signals obtained using the reference functions given in Eqs. (5.24) and (5.25). The procedure to determine the PDF for the ratio of two photon-limited correlation signals is given in Section 4.5.2, and can be applied here.

As before, we define the variables

$$Z \equiv \frac{\xi}{\eta}; \quad \xi \equiv |C_1|, \quad \eta \equiv |C_2| \quad . \quad (A17)$$

If different photoevents are used to realize the two correlation signals, then they will be statistically independent. The individual density functions $P(\zeta)$ and $P(\eta)$ are given by Eq.(A15), with the appropriate statistical moments used for each variable. The probability density function of the ratio of the two independent random variables is expressed in terms of their individual density functions via the following equation⁵:

$$P(Z) = \int_0^{\infty} \eta P_{\zeta}(Z\eta) P_{\eta}(\eta) d\eta \quad . \quad (A18)$$

In Eq. (A18), $P_{\zeta}(Z\eta)$ is obtained by substituting $Z\eta$ for ζ in the expression for $P(\zeta)$:

$$P_{\zeta}(Z\eta) \equiv \frac{Z\eta}{2\pi(1-\rho_{\zeta}^2)^{\frac{1}{2}}\sigma_{\zeta}^2} \exp\left\{-\frac{(Z\eta)^2 + \langle \zeta' \rangle^2 + \langle \zeta'' \rangle^2}{2(1-\rho_{\zeta}^2)\sigma_{\zeta}^2}\right\} I_0\left\{\frac{Z\eta(\langle \zeta' \rangle^2 + \langle \zeta'' \rangle^2)^{\frac{1}{2}}}{(1-\rho_{\zeta}^2)\sigma_{\zeta}^2}\right\} \quad . \quad (A19)$$

In Eq.(A19), η' and η'' denote $\text{Re}\{C_1\}$ and $\text{Im}\{C_1\}$ respectively. The statistical moments are obtained using Eqs.(A6)-(A11). In addition, $P_{\eta}(\eta)$ denotes the probability density function for η , and is given by

$$P_{\eta}(\eta) \equiv \frac{\eta}{2\pi(1-\rho_{\eta}^2)^{\frac{1}{2}}\sigma_{\eta}^2} \exp\left\{-\frac{(\eta)^2 + \langle \eta' \rangle^2 + \langle \eta'' \rangle^2}{2(1-\rho_{\eta}^2)\sigma_{\eta}^2}\right\} I_0\left\{\frac{\eta(\langle \eta' \rangle^2 + \langle \eta'' \rangle^2)^{\frac{1}{2}}}{(1-\rho_{\eta}^2)\sigma_{\eta}^2}\right\} \quad (A20)$$

Substituting Eqs.(A19) and (A20) into (A18), it is possible to obtain an analytic solution to the integration, using Ref. 6..The result of the integration is

$$\left[\frac{Z^2}{Z(1-\rho_{\xi}^2)^2\sigma_{\xi}^2} + \frac{1}{Z(1-\rho_{\eta}^2)^2\sigma_{\eta}^2} \right]^{-2} \sum_{i=0}^{\infty} \frac{(-1)^m \Gamma(m+2)}{m! \Gamma(m+1)} \left[\frac{\frac{Z^2(\langle \xi' \rangle^2 + \langle \xi'' \rangle^2)}{(\sigma_{\xi}^2(1-\rho_{\xi}^2)^2)^2}}{\frac{4Z^2}{Z(1-\rho_{\xi}^2)^2\sigma_{\xi}^2} + \frac{4}{Z(1-\rho_{\eta}^2)^2\sigma_{\eta}^2}} \right]^m \cdot {}_2F_1\left(-m, -m, 1, \frac{\frac{Z^2(\langle \eta' \rangle^2 + \langle \eta'' \rangle^2)}{(\sigma_{\eta}^2(1-\rho_{\eta}^2)^2)^2}}{\frac{Z^2(\langle \xi' \rangle^2 + \langle \xi'' \rangle^2)}{(\sigma_{\xi}^2(1-\rho_{\xi}^2)^2)^2}}\right) \quad (A.21)$$

In Eq. (A.21), $\Gamma()$ denotes the Gamma⁷, or factorial function, and $F()$ denotes the hypergeometric function⁸. It was the author's experience that the integration was more conveniently performed numerically, rather than evaluate the infinite series of hypergeometric functions. Once the PDF for the ratio of the correlation signals has been computed (numerically), the mean value and variance can also be computed numerically.

A.2. Statistics of the distance in feature space, D^2

The density function for the distance in feature space is needed to predict exactly the number of detected photoevents necessary to discriminate among a given set of images. Here, we present the procedure to determine the statistics of the distance in

feature space when $\Phi_{2,m}$ are used as reference features. A similar procedure is used when $\Phi_{1,m}$ are the reference features.

The distance in feature space D^2 , given by (c.f. Eq.(7))

$$D^2 = \sum_m [\Phi_{2,m}^{\text{input}} - \Phi_{2,m}^{\text{ref.}}]^2, \quad (\text{A22})$$

is computed using the photon-limited estimate for Φ^{input} (recall that the estimate for Φ^{input} is given by $|C_{2,m}|$ when Eq.(5.13) is used as the reference function with $s=2$). If different photoevents are used to estimate the different order moments, each estimate is statistically independent. Hence, the density function $P(D^2)$ is given by the convolution of the m density functions³² $P((|C_{2,m}| - \Phi^{\text{ref.}})^2)$. The density function $P((|C_{2,m}| - \Phi^{\text{ref.}})^2)$ can be obtained using Eq. (A15) and the following change of variables. Let

$$\mu_{2,m} = (|C_{2,m}| - \Phi_{2,m}^{\text{ref.}})^2, \quad (\text{A23})$$

or

$$|C_{2,m}| = \pm \mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}}. \quad (\text{A24})$$

Making this substitution into Eq.(A15) and applying the Jacobian for the transformation, one finds the following result for the density function $P(\mu_{2,m})$:

$$\begin{aligned}
P(\mu_{2,m}) = & \left[\frac{(\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}}) \exp - \left\{ \frac{\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}} + \langle C' \rangle^2 + \langle C'' \rangle^2}{2\sigma^2(1-\rho^2)} \right\}}{2\sigma^2 \sqrt{(1-\rho^2)} \mu_{2,m}} \right. \\
& \times I_0 \left\{ \frac{(\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}}) (\langle C' \rangle^2 + \langle C'' \rangle^2)^{\frac{1}{2}}}{\sigma^2(1-\rho^2)} \right\} \\
& + \left[\frac{(-\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}}) \exp - \left\{ \frac{-\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}} + \langle C' \rangle^2 + \langle C'' \rangle^2}{2\sigma^2(1-\rho^2)} \right\}}{2\sigma^2 \sqrt{(1-\rho^2)} \mu_{2,m}} \right. \\
& \times I_0 \left\{ \frac{(\mu_{2,m}^{\frac{1}{2}} + \Phi_{2,m}^{\text{ref.}}) (\langle C' \rangle^2 + \langle C'' \rangle^2)^{\frac{1}{2}}}{\sigma^2(1-\rho^2)} \right\} \left. \right] \quad (A25)
\end{aligned}$$

Hence, when different photoevents are used to estimate each moment, the density function for $P(D^2)$ is given by

$$P(D^2) = P(\mu_{2,m_1}) \otimes P(\mu_{2,m_2}) \otimes P(\mu_{2,m_3}) \dots \quad (A26)$$

where \otimes represents convolution and m_i represent only the orders of the angular moments that are used to compute the distance vector (in practice, the total number of

moments required is usually ≤ 3). If the same photoevents are used to compute each order moment, then the random variables $\mu_{2,m}$ are not statistically independent. The approximation that $\mu_{2,m}$ are independent may still be useful, but this must be determined either experimentally or numerically on a case-by-case basis.

A.3 Statistics of the photon-limited estimate for the input centroid.

The photon-limited estimates for moment invariants are shift invariant only when the moments are computed about the centroid of the input image. In addition, the centroid of the input must be known within a given error to produce accurate values for the moments invariants; the error that can be tolerated will depend on the particular application in question. (For the experiments performed here, the centroid was required within 0.5 pixels using 256x256 input images.) In this appendix, we present the method for determining the number of detected photoevents required to determine the centroid of the input image within a given error.

The Cartesian coordinates of centroid of the input image can be determined by realizing the correlation signals C_x and C_y using the reference functions

$$R_x(x, y) = x \quad , \quad (A27)$$

$$R_y(x, y) = y \quad . \quad (A28)$$

Upon consecutive substitutions of Eqs. (A27) and (A28) into (A6), one finds

$$\langle C_x \rangle = \frac{\iint_A x f(x, y) dx dy}{\iint_A f(x, y) dx dy} \quad , \quad (A29)$$

$$\langle C_y \rangle = \frac{\iint_A y f(x, y) dx dy}{\iint_A f(x, y) dx dy} , \quad (A30)$$

where A represents the area of the input image. The variance in the estimate for the x and y coordinates of the centroid is obtained by consecutively substituting Eqs. (A27) and (A28) into Eq. (A8). Hence,

$$\sigma_x^2 = \frac{1}{N} \left\{ \frac{\iint_A x^2 f(x, y) dx dy}{\iint_A f(x, y) dx dy} - \langle C_x \rangle^2 \right\} , \quad (A31)$$

$$\sigma_y^2 = \frac{1}{N} \left\{ \frac{\iint_A y^2 f(x, y) dx dy}{\iint_A f(x, y) dx dy} - \langle C_y \rangle^2 \right\} , \quad (A32)$$

Different applications will require that the centroid be computed with different errors; one definition for an error is given in Eq.(5.24). The amount of error that can be tolerated in the computation of the centroid must be established on a case by case basis, and Eqs. (A31) and (A32) can be used to determine the number of detected photoevents needed to achieve the desired error.

References for Appendix A

1. L. Mandel, E. C. G. Sudarshan and E. Wolf, 'Theory of photoelectric detection of light fluctuations,' *Proc. Phys. Soc.* **84**, 435-444 (1964).
2. M. Bertolotti, 'Photon Statistics,' in *Photon Correlation and Light Beating Spectroscopy*, H. Z. Cummins and E. L. Pike, eds. (Plenum, New York, 1974), Chap. 2.
3. I. S. Gradshteyn and Ryzhik eds., *Table of Integrals, Series, and Products* (Academic, New York 1980). p. 488.
4. Abramowitz and Stegun eds., *Handbook of Mathematical Functions* (Dover, New York, 1972) p. 376, Eq. (9.6.16)
5. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1965), p:197.
6. Ref. 3, p. 718
7. G. Arfken, *Mathematical Methods for Physicists*, (Academic, New York, 1970) p.450.
8. *Ibid.* p. 632-636.

Appendix B. Issues in the computation of weak moments

As mentioned in Section 5.5.3, it is not practical to compute weak radial moments using photon-counting techniques. The advantage in using photon-counting techniques is that one can often estimate radial moments of CHF's using a single realization of the photon-limited correlation signal, when the proper reference function is used. For this estimate to be accurate, the mean value of the estimate must be approximately equal to the actual value of the radial moment, and the standard deviation of the estimate must be small relative to the mean value of the estimate. Here, we show that this is indeed true for strong (or large) radial moments, but is not true for weak (or small) radial moments.

From Eq.(5.21), we see that the invariant moment $\Phi_{s,m}$ is given exactly by taking the ratio $|<C_{s,m}>|/|<C_{s,0}>|$. An estimate for $\Phi_{s,m}$ can be obtained by taking the ratio $|C_{s,m}|/|C_{s,0}|$, where $C_{s,m}$ and $C_{s,0}$ are obtained using Eq.(5.12), with the reference function given in Eq.(5.19) (using appropriate values for s and m). If different photoevents are used to compute $|C_{s,m}|$ and $|C_{s,0}|$, then

$$\left\langle \frac{|C_{s,m}|}{|C_{s,0}|} \right\rangle = \frac{\langle |C_{s,m}| \rangle}{\langle |C_{s,0}| \rangle} = \Phi_{s,m}^+ \quad , \quad (B1)$$

where $\Phi_{s,m}^+$ is the photon-limited estimate for the invariant moment $\Phi_{s,m}$. Comparing Eq. (B1) with Eq.(5.21), we see that $\Phi_{s,m}^+$ will be approximately equal to $\Phi_{s,m}$ if

$$\langle |C_{s,m}| \rangle \gg |<C_{s,m}>| \quad , \quad (B2)$$

and

$$\langle |C_{s,0}| \rangle \cong \langle C_{s,0} \rangle . \quad (B3)$$

The mean value of $|C_{s,m}|$ and $|C_{s,0}|$ can be obtained numerically using the PDF given in Eq.(A15). If this is performed for radial moments of arbitrary strength, we see that Eq.(B2) is not consistently true (see Tables 5.1-5.3).

We can see this effect analytically by examining whether

$$\langle |C_{s,m}|^2 \rangle \cong \langle C_{s,m} \rangle^2 . \quad (B4)$$

and

$$\langle |C_{s,0}|^2 \rangle \cong \langle C_{s,0} \rangle^2 . \quad (B5)$$

We examine Eqs. (B4) and (B5) because they illustrate the point more readily than the more complex analytical expressions for $\langle |C_{s,m}| \rangle$; in addition, $\Phi_{s,m}^2$ is also a valid definition of a moment invariant. Using Eq.(5.12), the mean value of $|C_{s,m}|^2$ is given by

$$\langle |C_{s,m}|^2 \rangle = \frac{1}{N^2} \left\langle \sum_{i=1}^N \sum_{j=1}^N R(r_i, \theta_i) R^*(r_j, \theta_j) \right\rangle , \quad (B6)$$

where the centroid of the input image is defined to be the origin, and * denotes complex conjugation. Rearranging terms in the summations, one finds

$$\langle |C_{s,m}|^2 \rangle = \frac{1}{N^2} \left\langle \sum_{i \neq j}^N \sum_{j=1}^N R_i R_j^* + \sum_{k=1}^N R_k R_k^* \right\rangle , \quad (B7)$$

where R_i, R_j and R_k denote $R(r_i, \theta_i)$, $R(r_j, \theta_j)$ and $R(r_k, \theta_k)$ respectively.

Computing the ensemble average term by term, one obtains

$$\begin{aligned}
\langle |C_{s,m}|^2 \rangle = & \frac{(N^2 - N) \int_0^{2\pi} \int_0^\infty f(r', \theta') R(r', \theta') r' dr' d\theta' \int_0^{2\pi} \int_0^\infty f(r', \theta') R^*(r', \theta') r' dr' d\theta'}{N^2 \left[\int_0^{2\pi} \int_0^\infty f(r', \theta') r' dr' d\theta' \right]^2} \\
& + \frac{1}{N} \frac{\int_0^{2\pi} \int_0^\infty f(r', \theta') |R(r', \theta')|^2 r' dr' d\theta'}{\int_0^{2\pi} \int_0^\infty f(r', \theta') r' dr' d\theta'} .
\end{aligned} \tag{B8}$$

In Eqs. (B6)-(B8), the reference functions are given by Eq.(5.19), using the appropriate values of s and m . For N large and $f(r, \theta)$ real, the first term in Eq.(B8) is approximately $|M_{s,m}|^2/|M_{2,0}|^2$, [cf. Eq.(5.3)]; this is the term that we hope is dominant. If the moment invariant $|M_{s,m}|^2/|M_{2,0}|^2$ in question is a sufficiently strong moment for the image in question, the second term in Eq.(B8) can be neglected with respect to the first term. As a result, Eq. (B4) will hold. The same is then true for Eq.(B5). In the particular case of $|C_{2,m}|^2$, Eq.(B8) becomes

$$\begin{aligned}
\langle |C_{2,m}|^2 \rangle = & \frac{\int_0^{2\pi} \int_0^\infty f(r', \theta') \exp\{-im\theta\} r' dr' d\theta' \int_0^{2\pi} \int_0^\infty f(r', \theta') \exp\{+im\theta\} r' dr' d\theta'}{\left[\int_0^{2\pi} \int_0^\infty f(r', \theta') r' dr' d\theta' \right]^2} \\
& + \frac{1}{N} .
\end{aligned} \tag{B9}$$

Hence, for N large,

$$\langle |C_{2,m}|^2 \rangle \cong \Phi_{2,m}^2 , \tag{B10}$$

and the mean value of the estimate for $\Phi_{2,m}^2$ is approximately equal to $\Phi_{2,m}^2$. While it does not directly follow that this is true for $\Phi_{2,m}$, one can make a similar argument by computing $\langle |C_{2,m}| \rangle$ numerically using the PDF in Eq.(A15). The results of this computation are given in Tables 5.1-5.3.

The variance for the estimate of $\Phi_{2,m}$ is given by

$$\sigma_{2,m}^2 = \langle \Phi_{2,m}^2 \rangle - \langle \Phi_{2,m} \rangle^2 \quad (B11)$$

Using Eqs. (B9)-(B11), we see that the variance in the estimate for $\Phi_{2,m}$ is of the order $1/N$, which is typically small compared to $\Phi_{2,m}$ if the moment in question is large (see sigma in Table 5.2, $m = 2$ or 4).

For the case when $1/N$ is not small compared to the first term of Eq. (B9), the effect of this bias can be reduced by using

$$\left(|C_{2,m}|^2 - \frac{1}{N} \right)^{\frac{1}{2}} \quad (B12)$$

as the estimate for the moment invariant. The PDF for this estimate of $\Phi_{2,m}$ can be obtained using Eq.(A15), and making the following change of variables. Let

$$\beta = \left(|C_{2,m}|^2 - \frac{1}{N} \right)^{\frac{1}{2}} \quad (B13)$$

or

$$|C_{2,m}| = \left(\beta^2 + \frac{1}{N} \right)^{\frac{1}{2}} \quad (B14)$$

Substituting Eq. (B14) into Eq.(A15), and applying the appropriate Jacobian of the transformation, one finds for the PDF $P(\beta)$

$$P(\beta) = \beta \exp \left\{ -\frac{\beta^2 + \frac{1}{N} + \langle C' \rangle^2 + \langle C'' \rangle^2}{2(1-\rho^2)\sigma^2} \right\} I_0 \left\{ \frac{(\beta^2 + \frac{1}{N})^{\frac{1}{2}}}{(1-\rho^2)\sigma^2} (\langle C' \rangle^2 + \langle C'' \rangle^2)^{\frac{1}{2}} \right\} . \quad (B15)$$

Eq.(B15) can be used in place of Eq.(A15) in all calculations regarding the statistical behavior of the photon-limited estimate of $\Phi_{2,m}$ when Eq.(B13) is used to provide the estimate.